International Encyclopedia of Marketing

perception of brand equity

Randle D. Raggio

What Brand Equity is

Derived from the concept that a brand represents a promise of salient benefits to a set of target consumers (Blackett, 2003), *brand equity* is the perception that a brand delivers on its promise of benefits (Raggio and Leone, 2007). This definition implies that brand equity is an intraindividual concept – that is, it lives in the hearts and minds of consumers. It is also context specific, in that a person can hold a high amount of brand equity for a brand on particular occasions, but not on other occasions, as the salience and importance of the promise, and the brand's ability to deliver on that promise, can be perceived to change across occasions (e.g., Lava soap may enjoy high brand equity among do-it-yourself auto mechanics in the shop, but this equity may not transfer to more common domestic cleaning situations) (Raggio and Leone, 2007).

What Brand Equity Does

Brand equity moderates the impact of brand-related information such as advertising, product-use experience, and word of mouth, producing more favorable customer reactions toward a product and the way it is marketed when the brand is identified, as compared to when it is not (Keller, 2008). Those who hold high amounts of brand equity are more likely to believe positive information about the brand, less likely to believe, or more likely to discount or counterargue, negative information, and more likely to give the brand the benefit of the doubt (e.g., when faced with product quality problems, product harm crises, or price or product challenges from competitive brands, etc.).

Outcomes of Brand Equity

Potential outcomes of brand equity include positive brand image, higher perceived quality, brand-name awareness, brand loyalty, brand choice, higher willingness to pay a price premium, more favorable reaction to brand extensions, less vulnerability to competitive marketing tactics, and so on (Keller, 2008), and it may produce these effects in consumers, channel partners, potential employees, government/regulatory agencies, financial institutions, and so on (Raggio and Leone, 2007). As a result, brand equity is a valuable intangible asset for a firm, through its ability to positively impact the cash flow of the firm that owns it or uses it, for example, through licensing agreements (*see* BRAND VALUE for more information on how brand equity impacts brand value). However, related to the idiosyncratic nature of brand equity described above, some consumers may hold high levels of brand equity but not (yet) be prospective users of the brand, so the lack of specific marketplace behaviors by a particular set of consumers (e.g., purchase) is not evidence that brand equity does not exist (e.g., luxury goods, baby-care products, college education, etc.) (Raggio and Leone, 2007). Likewise, because brand equity is based on consumer perceptions, specific marketplace behavior (e.g., purchase – even at a price premium) is not necessarily an indication that brand equity does exist, as some purchases may be driven by product attributes or ingredients, price, induced behavioral loyalty (e.g., airline frequent-flyer membership), or availability (Raggio and Leone, 2007).

Brand Equity as an Asset

Though brand equity is a valuable intangible asset, it is not a separable asset that can be traded in the marketplace like a patent, trademark, logo, or certain brand associations (e.g., paid endorsements) (Barwise, 2003), or valued accurately apart from the specific firm that owns or uses the brand, or a context-specific usage or consumer behavior. For example, it may be possible to determine the impact of brand equity on favorable impressions of a brand extension, but the effect may be different if the extension were produced by a license, or if the focal outcome were changed to willingness to pay a price premium. Measures of brand equity should consider both its immediate effects on current prospects and its ability to attract new users who are not currently prospects for what the brand has to offer (Raggio and Leone, 2007).

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Building Brand Equity

Brand equity is built by what the brand *has* and what the brand *does* to create the perception that the brand indeed meets its promise of benefits. Strong, positive, and unique brand associations are necessary to build brand equity, and can be driven by the product itself through salient attributes/features/ingredients and through its performance, linkages to positive feelings, memories, images, judgments, and so on, or by other proprietary brand assets such as patents, trademarks, exclusive distribution contracts, or endorsements that competitors cannot copy, to create a brand that resonates with consumers in context-specific usage (Keller, 2008).

Bibliography

Barwise, P. (2003) Preface, in *Brands and Branding*, (eds R. Clifton and J. Simmons), Bloomberg Press, Princeton, pp. xii–xv.

Blackett, T. (2003) What is a brand? in *Brands and Branding*, (eds R. Clifton and J. Simmons), Bloomberg Press, Princeton, pp. 13–15.

Keller, K.L. (2008) *Strategic Brand Management: Building, Measuring and Managing Brand Equity*, 3rd edn, Pearson Education, Upper Saddle River.

Raggio, R.D. and Leone, R.P. (2007) The theoretical separation of brand equity and brand value: managerial implications for strategic planning. *Journal of Brand Management*, **14**, 380–395.

customer equity

Katherine N. Lemon and Loren J. Lemon

Customer equity is defined as the total lifetime value of a firm's current and potential customer base. The customer equity (CE) construct can be utilized via both a conceptual framework (CE framework) and a mathematical model (CE model) which will allow a firm to evaluate various marketing initiatives, either conceptually or by a potential-ROI analysis, in order to create, manage, and grow the firm's CE (Blattberg and Deighton, 1996; Rust, Lemon, and Zeithaml, 2004; Rust, Zeithaml, and Lemon, 2000). The CE framework utilizes three primary elements that drive the creation and dynamic growth of CE: value, brand, and relationship.

The "value driver" is the customer's assessment of the utility of the firm's offering based on the customer's objective perception of what is given up for what is received. Typical subcomponents are *quality*, *price*, and *convenience*. *Quality* encompasses the customer's objective perception of the physical and nonphysical attributes of the offering. *Price* represents "what is given up by the customer," usually, but not always, money. *Convenience* relates to the customer's time and other costs in obtaining the offering and doing business with the firm (*see* SERVICES MARKETING STRATEGY; A FRAMEWORK FOR CREATING VALUE PROPOSITIONS).

The "brand driver" is the customer's subjective/intangible assessment of the firm and the offering. Typical subcomponents are *brand awareness*, *brand perceptions*, and *firm ethics*. *Brand awareness* encompasses those firm actions (e.g., advertising) that influence and enhance the customer's awareness of the subject brand. *Brand perceptions* deals with the customer's attitude toward that brand (i.e., how much the customer identifies with, or subjectively values, the brand). *Firm ethics* deals with the firm's acts or omissions that influence the customer's subjective perceptions of the firm (e.g., charitable/community activities, privacy policies, and corporate misconduct) (*see* PERCEPTION OF BRAND EQUITY; BRAND VALUE; BRAND GROWTH STRATEGY; BRAND STRATEGY).

The "relationship driver" is the customer's tendency to stick with a firm for reasons other than just the customer's objective and subjective assessments of the firm's offerings or brand, and focuses on the direct (or indirect) interactions between the customer and the firm. Typical subcomponents are *loyalty programs*, *affinity programs*, and *community-building programs*. *Loyalty programs* reward the customer for specific behaviors (e.g., flying on the airline) with tangible benefits (e.g., free tickets). *Affinity programs* seek to create strong emotional ties with the customer by connecting the firm's offering with other aspects of the customer's life (e.g., college alumni Visa cards). *Community-building programs* link a customer with a group of other like customers of the firm (e.g., Harley Owner Groups (HOGS)) (*see* SERVICES MARKETING STRATEGY).

Implementation of the CE framework begins with the determination regarding which of the drivers are key for the target customers. (This can be determined by a variety of ways; usually target customers are surveyed to determine their perceptions of the company's and its competitors' performance with regard to the CE drivers) (*see* CUSTOMER ANALYSIS). The firm then assesses its performance on the identified key drivers, and compares its performance with the relative importance of such drivers to the customer and relative to its competitors.

Using additional data necessary to calculate individual average CUSTOMER LIFETIME VALUE (CLV) and the aggregate firm CE, the firm can utilize the CE model to determine the potential relative impact that various marketing initiatives may have on the firm's CE. (While the detailed mathematics behind the model are beyond the scope of this article, the model takes into account brand switching patterns, Markov chain analysis, demographic projections, life expectancy forecasts, and value discounting) (*see* CUSTOMER RELATIONSHIP MANAGEMENT). By running the CE model on a variety of potential marketing initiatives, the firm can then compare and rank the various initiatives on a "CE return on marketing investment" perspective.

With insights derived from the CE framework and CE model, a firm can tailor its marketing

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

efforts to meet unfulfilled customer expectations, and identify and exploit situations in which the firm has a competitive advantage vis-à-vis its competitors (*see* MARKETING STRATEGY MODELS; A FRAMEWORK FOR CREATING VALUE PROPOSITIONS). This allows the firm to make a more disciplined and informed decision on which marketing initiatives to implement and at what levels.

Bibliography

Blattberg, R. and Deighton, J. (1996) Manage marketing by the customer equity test. *Harvard Business Review*, **74** (4), 136–144.

Rust, R.T., Lemon, K.N., and Zeithaml, V.A. (2004) Return on marketing: using customer equity to focus marketing strategy. *Journal of Marketing*, **68**, 109–127.

Rust, R.T., Zeithaml, V.A., and Lemon, K.N. (2000) *Driving Customer Equity: How Customer Lifetime Value is Reshaping Corporate Strategy*, The Free Press, New York.

brand growth strategy

Victoria L. Crittenden

INTRODUCTION

The strategic opportunity matrix identifies four major strategic alternatives for company growth: market penetration, market development, product development, and diversification (Ansoff, 1957). Brands are fundamental components of each of these strategic growth options, and the company's MARKETING STRATEGY depends strongly on decisions surrounding brand strategies (*see* BRAND STRATEGY) within the larger context of MARKETING PLANNING.

Evolving over time, brand strategies range from corporate brands to individual brands. This evolution is driven by competitor innovations, changes in customers' wants and needs, and overall neglect that leads to atrophy of brand value (Capon *et al.*, 2001). The potential harm to BRAND VALUE, brand equity (*see* PERCEPTION OF BRAND EQUITY), and CUSTOMER EQUITY has long-term ramifications for shareholder value. Both business-to-consumer and business-to-business companies should plan strategically with respect to brand architecture. That is, all companies must understand the various strategies for brand growth so as to create brand value/equity and customer equity which lead to long-term shareholder value.

Brand growth strategies. The five major strategies for brand growth are line extension strategy, brand extension strategy, cobranding strategy, flanker strategy, and new brand strategy. There are no hard and fast rules as to when one brand growth strategy takes precedence over another. Brand growth requires a unique blend of art and science. Careful monitoring of the marketplace with respect to competition and customer desires is critical to brand growth strategy formulation.

Line extension strategy. The line extension strategy, probably the most common brand growth strategy, is a powerful tool in a product manager's toolbox. A line extension is when a parent brand is used to brand a new product that targets a new market segment within the product category served currently by the parent brand (Keller, 2007). A fairly common example of the line extension strategy is Crest toothpaste. A quick scan of the Procter and Gamble website highlights the depth of line extensions. For example, Crest toothpaste comes in a wide variety of product offerings: Pro Health, Weekly Clean, Whitening Expressions, Whitening, Cavity Protection, Tartar Protection, Sensitive Teeth, Flavors, Baking Soda, Gels, Liquid Gels, Paste, Striped, and For Kids.

In the Crest example, the relationship between the core brand (Crest) and the line extensions is that Crest serves as the endorser of the extensions. A consumer may purchase whitening toothpaste for the benefits of whitening, but purchase Crest Whitening for the assurance of the Crest brand. Yet, Whitening would likely carry no weight as a stand-alone brand without the endorsement of Crest. In other instances of line extension, both the core brand and line extension exert influence on consumer decision making. The Toyota Avalon, Toyota Camry, and Toyota Corolla are examples of line extensions where the core brand (Toyota) and the extensions (Avalon, Camry, and Corolla) are important designators of brand building power.

Product managers must keep in mind the line logic when considering line extensions. Line Logic™ is a disciplined approach to creating and presenting a full complement of product line extensions (Reffue and Crittenden, 2006). The six interrelated elements in the Line Logic™ framework are segmentation, naming conventions, design elements, tiered feature sets, pricing, and packaging. In considering the logic behind line extensions, this framework is useful for building a strong portfolio of line extensions – all of which work to complement each other and increase the overall market share and profitability. Line extensions should be a source of new revenue for a company; yet, one of the biggest fears is that the line extension will fail and damage the core brand in its failure.

Brand extension strategy. While a line extension uses the core brand name in the same product category as the original core brand, a brand extension strategy uses the brand name asset to penetrate new product categories with a new product (Aaker, 1990). The brand extension strategy utilizes brand name recognition as a way to have immediate name recognition in

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

a particular product category. Seven types of brand extensions have been suggested (Tauber, 1988). These seven types and examples of each are as follows:

1. Same product in a different form – this is one of the simplest forms of brand extension; an example of this form is when Ocean Spray created cranberry juice cocktail (a drink) by using its original food product (cranberries).
2. Distinctive taste/ingredient/component in new item – this is when a company takes a key ingredient or component of a branded product and uses that component in another product; an example is Arm & Hammer (original product was baking soda) with its Carpet Deodorizer.
3. Companion products – this is the extension of the brand into a product that can be used with the original branded product, such as Log Cabin syrup offering Log Cabin pancake mix.
4. Same customer franchise – an attempt to leverage a brand name by selling something additional to the current customer base; an example would be the American Express gift card, which is a typical gift card but leverages the American Express brand and reputation.
5. Expertise – this is when a brand extension is offered in a product category for which the customer expects the company to have expertise; for example, BIC was the leader in disposable plastic pens and transferred this disposability expertise to disposable lighters and razors.
6. Benefit/attribute/feature owned – this is when consumers tend to associate a particular attribute with a brand; an example of this is when Ivory soap used its mildness attribute to extend into a mildness positioning for shampoo for daily use.
7. Designer image status – this is when the status of the original brand extends its status to other product categories, such as the Porsche name appearing on sunglasses.

A brand extension strategy can help combat entry barriers put in place by prominent brand name domination of particular product categories. Consistent with this, the primary benefits to the brand extension strategy are capitalization on the company's most valuable asset – its brand names, investment outlays typically necessary to establish a new brand are minimal, introduction of the new brand can increase sales for the parent brand, and reduced risk of failure (Tauber, 1988).

Yet, there can be a dark side to brand extensions. Failures are attributed to a variety of reasons (Aaker and Keller, 1990; Capon *et al.*, 2001). For example, a lack of association between old and new products can lead to consumer confusion or even consumer disinterest. Oftentimes, the unique image of the original brand is not transferable to the brand extension. Even worse, the extension might be perceived to be of lower quality than the original product with the same brand name. Competitively, there might already be a dominant brand in the new product category, which could lead to unexpected and/or time-consuming competitive battles.

Cobranding strategy. The cobranding growth strategy is the formation of an alliance between popular brands. While not necessarily new to the marketplace, this cobranding growth strategy has become more prevalent given the popularity of alliances as a growth strategy in general. The synergy created by two well-matched brands can clearly enhance long-term value to the firm, with products signaling unobservable qualities by having a brand ally (Blackett and Boad, 1999; Rao, Qu, and Ruekert, 1999).

A physical cobranding growth strategy occurs when the two brands come together to form one product. Examples include a Dell PC with Intel inside and Diet Coke with NutraSweet. Ford Motor Company, however, has tried its hand at a different form of a cobranding growth strategy. In the mid-2000s, the company introduced the Ford F-150 Harley–Davidson. In the early 1990s, Ford cobranded with Eddie Bauer to offer Eddied Bauer branded Ford vehicles. Both of these Ford examples represent a cobranding that is more symbolic in nature in that neither Harley–Davidson nor Eddie Bauer provided component parts to the vehicles. Instead, the cobranding strategies at Ford were meant to relate Ford and Harley–Davidson/Eddie Bauer in the minds of the consumers and

signal an unobservable aspect of the Ford vehicles.

Following a cobranding growth strategy has many benefits, not the least of which is bringing together the power of two strong brands to the marketplace thus enabling a new product to tap into the market following of both brands. On the flip side, however, companies must be careful when bringing together disparate brands that might cause confusion in the consumer's mind. For example, some consumers might not opt for the Eddie Bauer version of the Ford if they are unsure as to the significance of an Eddie Bauer vehicle, instead attributing Eddie Bauer to a line of clothing or retail stores. Additionally, neither company has control over the actions of the cobranding partner. Thus, something that harms the reputation of one could spill over and have a negative impact on the cobranded product.

Flanker strategy. The flanking strategy has its roots in warfare. As related to marketing, the flanking growth strategy is one of the MARKETING WARFARE STRATEGIES. The general idea is that a company positions products in such a way as to have a presence in various market segments. In military jargon, the flanking growth strategy reduces the maneuverability of the enemy (competitor), with a defensive flanking strategy referred to as a *flanking position* and the offensive strategy referred to as a *flanking attack*. As a defensive move, a company following a flanking growth strategy will offer products in peripheral/secondary markets so as to prevent competitive attacks on a weak or no brand position. Offensively, the flanking growth strategy is engaged to avoid a head-on or direct confrontation with an established competitor.

The flanking brand growth strategy suggests that the brand is a fighting brand. A company following this strategy generally does not want to use its leading brand in battle, as the risks are too great. Thus, the company brings out the fighting brand to battle on possibly less strategic fronts and/or to prevent the competition from gaining market entry from a vulnerable position.

Cytosport, the parent company of Muscle Milk, understands the flanking brand growth strategy. With Muscle Milk as a market leader in a growing health and nutrition marketplace, Cytosport recognized that other companies would want a piece of this rapidly expanding marketplace. Rather than only attempting to capture this growth via line or brand extensions, Cytosport focused on Cytomax as its flanking brand. Cytomax was intended to head off attacks in expanding market segments. The company's strategy has been to offer a core brand and then have strategically derived flanking brands in market segments as necessary.

Pharmaceutical companies have also begun using flanking strategies by aligning with generic competitors in an effort to capture a share of the generic marketplace. Use of the flanking strategy in this context means that large pharmaceutical companies are not fighting head to head with the generic firms; rather the large pharmaceutical companies are using a form of cobranding as a flanking brand strategy.

A major risk in the flanker brand growth strategy is that the market segments may not be as clearly delineated as hoped. That is, consumers might trade down to the flanker brand. Thus, this is a particularly risky brand growth strategy in difficult economic times.

New brand strategy. The new brand growth strategy leverages the existing brand's marketing clout in the introduction of a completely new brand. The company may be attempting to tap into new markets that might not fit the profile of the current customer base. Thus, there is a new brand offering in a product or market class that was not served previously by the company.

Toyota has followed the new brand growth strategy quite successfully with its Lexus and Scion product offerings. Toyota introduced the Lexus to compete in the very high end of the market – a consumer market that might have been turned off by a Toyota product which is considered more for the middle, mainstream marketplace. Similarly, in the mid-2000s, Toyota sought to capture the younger demographic with the Scion. While the Lexus consumer did not see the Toyota as an upscale automobile, the younger marketplace had a difficult time seeing Toyota as a trendy and fun car to drive. Thus, Toyota entered this market with the Scion. While both Lexus and Scion consumers know that Toyota is the

parent company, neither group of consumers see themselves as driving a Toyota automobile. Yet, both groups know that their vehicles are backed by Toyota quality even without the Toyota brand name on the product.

In a similar move, Pepsi expanded into the non-cola market with a tea beverage named Tava. A fruit-flavored, caffeine-free drink, Tava allowed Pepsi to enter a new market with a new product. Branding the product as a Pepsi product would have enforced a strong association with a cola drink – something Tava is not. Additionally, the entirely new brand allowed Pepsi to experiment with some non-traditional media alternatives without risking Pepsi's traditional branded products.

The Toyota and Pepsi examples portray new brand strategies that work for companies. In both instances, the companies were extending into comparable product categories in seeking new consumers. An example of a new brand strategy that did not work well was the effort by Anheuser–Busch to break into the salty snack market with Eagle Snacks. While eating salty snacks might increase beer consumption, a beer company offering a new brand of salty snacks did not resonate with consumers.

The new brand strategy enables a company to grow without compromising its reputation in already-strong markets. A company is able to exert influence in the marketplace with its marketing clout, but it is not putting its reputation at risk. Yet, since there is no guarantee that consumers will naturally connect the goodness of the core brand with the new brand, the company is also foregoing any possible positive spillover effect.

Timing of brand growth strategies. Prominent in marketing strategy is the idea of the first mover advantage (*see* FIRST-MOVER (PIONEER) ADVANTAGE), yet brand growth strategies do not necessarily follow this early or first entry idea. For example, one study found that brand extensions entering early in the market lifecycle were more likely to fail (Sullivan, 1992). Entering late in the market lifecycle was suggested as the preferred strategy for a brand extension. Within the same study, it was found that the new brand growth strategy was likely more effective in the early stages of the market lifecycle, with early

Table 1 Summary of brand growth strategies.

Strategy	*Core Brand*	*Growth*
Line extension	Crest (toothpaste)	Pro health, weekly clean, whitening expressions, whitening, cavity protection, tartar protection, sensitive teeth, flavors, baking soda, gels, liquid gels, paste, striped, for kids
	Toyoto (cars)	Avalon, Camry, Corolla
Brand extension	Ocean spray	Same product different form
	Arm & Hammer	Key component in new item
	Log Cabin	Companion product
	American Express	Same customer franchise
	BIC	Expertise
	Ivory	Benefit/attribute
	Porsche	Designer image status
Cobranding	Dell PC	Intel
	Diet Coke	NutraSweet
	Ford Motor Co.	Harley–Davidson/Eddie Bauer
Flanker	Muscle milk	Cytomax
	Pharmaceutical Cos.	Generic manufacturers
New brand	Toyota	Lexus &and Scion
	Pepsi	Tava

Table 2 Disadvantages of brand growth strategies.

Strategy	*Disadvantages*
Line extension	Damage to core brand; cannibalization; loss of product/brand focus; dilution of brand image; brand confusion; limited shelf space
Brand extension	Consumer confusion; consumer disinterest; lack of transferability; unequal quality; competitive battles; potential damage to core; intrafirm competition; cannibalization
Cobranding	Disparate, not complementary, brands; negative spill over effect reputation-wise; increased challenge of two brands; potential conflict between companies
Flanker	Spillover between market segments; increased resources; increased competition; cannibalization
New brand	Lack of brand connection; typical new product risks

entering new brand strategies outperforming late entering new brand strategies. Using the product lifecycle (*see* STAGES OF THE PRODUCT LIFE CYCLE) framework, companies tend to pursue brand growth strategies, particularly line extensions, in the growth stage of the product lifecycle.

Conclusion

Table 1 summarizes the five brand growth strategies with respect to core brands and implemented growth strategies. While the benefits of each brand growth strategy are evident in long-term value, Table 2 provides an overview of the possible disadvantages associated with each brand growth strategy. Unfortunately, there are no magic formulas for managing brand growth. Brand growth decisions are judgment calls utilizing both art and science. They are based on an analysis of market situations, company knowledge, and, of course, a hint of intuition. Not surprisingly, all five brand growth strategies are intertwined closely with the product and market lifecycles.

Bibliography

Aaker, D.A. (1990) Brand extensions: the good, the bad, and the ugly. *Sloan Management Review*, **31**, 47–56.

Aaker, D.A. (2004) *Brand Portfolio Strategy: Creating Relevance, Differentiation, Energy, Leverage, and Clarity*, The Free Press, New York.

Aaker, D.A. and Keller, K.L. (1990) Consumer response to brand extensions. *Journal of Marketing*, **54**, 27–41.

Ansoff, H.I. (1957) Strategies for diversification. *Harvard Business Review*, **35**, 113–124.

Blackett, T. and Boad, B. (1999) *Co-Branding: The Science of Alliance*, Macmillan Press, London.

Capon, N., Berthon, P.R., Hulbert, J.M., and Pitt, L.F. (2001) Brand custodianship: a new primer for senior managers. *European Management Journal*, **19**, 215–227.

Keller, K.L. (2007) *Strategic Brand Management*, Prentice Hall, Upper Saddle River.

Rao, A.R., Qu, L., and Ruekert, R.W. (1999) Signaling unobservable product quality through a brand ally. *Journal of Marketing Research*, **36**, 258–268.

Reffue, D. and Crittenden, V.L. (2006) Line logic™ on the bow tie. *Journal of Product and Brand Management*, **15**, 168–172.

Sullivan, M.W. (1992) Brand extensions: when to use them. *Management Science*, **38**, 793–806.

Tauber, E.M. (1988) Brand leverage: strategy for growth in a cost-control world. *Journal of Advertising Research*, **28**, 26–30.

brand strategy

Colin Jevons

Brand Origins

In the days when livestock roamed free, farmers identified their property by burning an indelible mark on their animals' skins, originally with a piece of burning wood. This practice was named after the Old Norse word for burning – "brandr." The first Western commercial brand was probably that of Bass beer in England in the seventeenth century, a triangle scorched on the surface of its barrels to certify its provenance from a good, quality-controlled source. But there is evidence of brands in China well before that, with an image of a white rabbit used as a symbol of quality for sewing needles as far back as the Song Dynasty, from 960 to 1127 CE (Eckhardt and Bengtsson, 2007). Just as agriculture has developed since the early days, starting with the development of separate enclosures for animals up to modern-day bar-coded ear tags and RFID, so too has the wider application of branding and brand strategy developed (particularly since the mid-twentieth century) and the concept of a brand being merely an identifying sign, symbol, or device for differentiation is now very outdated. Brand strategies that worked in the twentieth century are becoming less effective in the twenty-first as the nature of brands and how they are cocreated changes (*see* BRAND GROWTH STRATEGY). de Chernatony (2009) suggests that any definition of brand should include an appreciation of differentiation, positioning, personality, vision and added value in providing a cluster of values that enable a promise to be made about a unique and welcomed experience. In contrast, the *American Marketing Association* offers three definitions on its website: the first definition reads "A name, term, design, symbol, or any other feature that identifies one seller's good or service as distinct from those of other sellers. The legal term for brand is trademark. A brand may identify one item, a family of items, or all items of that seller. If used for the firm as a whole, the preferred term is trade name." While the second definition includes a reference to the experiential, "a customer experience represented by a collection of images and ideas," it focuses mainly on the symbolism of the product offering, and the third definition specifically refers to tangible representations such as logos which can be developed to represent values. Since strategic marketing managers devote more effort to customer-focused activities than to trademarks and logo design, this article will consider strategy as it relates to the customer experience of the brand, following Keller's (1993) principle of customer-based brand equity. This diversity of views shows that any contemporary understanding of "brand" must be highly contextualized (Gabbott and Jevons, 2009), adding complexity to any attempt to define brand strategy.

Brand Strategy

Given this complexity, it may be best to refer to an authority that is as free of contextual bias as possible. The Oxford English Dictionary (1989) reports definitions taken from public usage, and in this way defines strategy in a business sense as a plan for successful action (*see* MARKETING STRATEGY). Planning the success of brands involves a wide range of actions and therefore a wide range of managerial decisions (*see* MARKETING PLANNING). Brand strategists must make decisions about brand hierarchies (corporate, family, and individual brand), portfolios of brands, and consolidating and focusing brands in the transition from simple transactional branding toward brand relationships with consumers (*see* PORTFOLIO MANAGEMENT; CONSUMER INVOLVEMENT). This increased emphasis on relationships is leading to a revival of the conflict between private label and manufacturer brands. The dynamic business and societal environment is increasing the need for managers to be aware of the drivers of brand positioning and repositioning, maintaining a delicate balance between the need for a brand to deliver a consistent message to its stakeholders while still remaining relevant to them (Blumenthal, 2002). Astute brand strategists also must manage the tension between attempting to differentiate the brand while ensuring that it continues to deliver on the performance dimensions most salient

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

to buyers (Ehrenberg, Barnard, and Scriven, 1997). The relationship between brands and their brand extensions is also important, as the extension can affect the parent brand itself – either positively or negatively (*see* BRAND EXTENSIONS AND FLANKER BRANDS). A further challenge with brand strategy today is that managers no longer have complete control over brand positioning and brand meaning (if they ever did have it; *see* POSITIONING ANALYSIS AND STRATEGIES). Courtesy of new communications technologies and consumer empowerment, cocreated brand meaning is one of the more pressing challenges facing twenty-first century brand managers, along with consumer-led brand positioning (Beverland and Ewing, 2005; *see* VALUE CO-CREATION). Thus, brand strategists may at times need to "manage" multiple brand meanings, whereby the same brand may come to mean different things to different stakeholders despite seemingly consistent positioning and marketing communications messages (Berthon, Pitt, and Campbell, 2009). Additional challenges facing brand strategists in the short-to-medium term include managing consumer backlash, anticonsumption, and the rejection of brand hegemony (Cromie and Ewing, 2009), recognizing and managing the inevitability of brand demise: that is, not all brands can nor should be saved (Ewing, Jevons, and Khalil, 2009), and developing more effective approaches to valuing brand equity and brand capability (Ratnatunga and Ewing, 2009, *see* BRAND VALUE). Finally, the extant literature on brand strategy has a distinctly western, distinctly blue-chip bias. There is still much we need to learn about eastern, for example, Asian, generally (Ewing, Napoli, and Pitt, 2001) Chinese (Ewing *et al.*, 2002) and Indian, approaches to brand management. Similarly, while much has been written about managing Fortune 500 and Interbrand top 100 (largely iconic) brands, less is written about managing brands within small and medium enterprises (Berthon, Ewing, and Napoli, 2008). Thus, despite a reasonably impressive corpus of brand strategy knowledge, there is still a fair amount more for brand strategists and scholars to learn.

Bibliography

Simpson, J., and Weiner, E. (1989) *Oxford English Dictionary*, 2nd edition, Oxford University Press, Oxford.

American Marketing Association (2009) Definition of Brand, http://www.marketingpower.com/_layouts/Dictionary.aspx?dLetter=B (accessed 2 October 2009).

Berthon, P., Ewing, M.T., and Napoli, J. (2008) Brand management in small-to-medium sized enterprises. *Journal of Small Business Management*, **46** (1), 27–45.

Berthon, P., Pitt, L., and Campbell, C. (2009) Does brand meaning exist in similarity of singularity? *Journal of Business Research*, **62** (3), 356–361.

Beverland, M. and Ewing, M.T. (2005) Slowing the adoption and diffusion process to enhance brand repositioning: the consumer driven repositioning of Dunlop Volley. *Business Horizons*, **48**, 385–391.

Blumenthal, D. (2002) Beyond 'from versus content': simmelian theory as a framework for adaptive brand strategy. *Journal of Brand Management*, **10** (1), 9–18.

de Chernatony, L. (2009) Towards the holy grail of defining 'brand'. *Marketing Theory*, **9** (1), 101–105.

Cromie, J.G. and Ewing, M.T. (2009) The rejection of brand hegemony. *Journal of Business Research*, **62** (1), 218–230.

Eckhardt, G. and Bengtsson, A. (2007) Pulling the white rabbit out of the hat: consuming brands in imperial China. Paper presented at the Association for Consumer Research European Conference in Milan, Italy.

Ehrenberg, A., Barnard, N. and Scriven, J. (1997) Differentiation or Salience. *Journal of Advertising Research*, **37** (6), 7–14.

Ewing, M.T., Jevons, C.J., and Khalil, E.J. (2009) Brand death: a developmental model of brand senescence. *Journal of Business Research*, **62** (3), 332–338.

Ewing, M.T., Napoli, J., and Pitt, L.F. (2001) Managing southeast Asian brands in the global economy. *Business Horizons*, **43** (2), 52–58.

Ewing, M.T., Napoli, J., Pitt, L.F., and Watts, A. (2002) On the renaissance of domestic brands in China. *International Journal of Advertising*, **21** (2), 197–216.

Gabbott, M. and Jevons, C. (2009) Brand community in search of theory: an endless spiral of ambiguity. *Marketing Theory*, **9** (1), 119–122.

Keller, K.L. (1993) Conceptualising, measuring, and managing customer-based brand equity. *Journal of Marketing*, **57**, 1–22.

Ratnatunga, J. and Ewing, M.T. (2009) An ex-ante approach to brand capability valuation. *Journal of Business Research*, **62** (3), 323–331.

brand value

Randle D. Raggio

What Brand Value is?

Brand value is the sale or replacement price of a brand (Raggio and Leone, 2009). Although brands are valuable to consumers (by reducing search costs and risk) (Barwise, 2003), the concept of brand value most frequently is applied from the perspective of a business that owns, potentially owns, or uses a brand (*see* PERCEPTION OF BRAND EQUITY for more information about consumer impacts of brands). Brands create value for businesses by positively impacting cash flows through increased revenue (e.g., higher probability of choice, insulation from competitive threat, better acceptance of brand extensions, etc.) or reduced costs (e.g., the ability of certain brands to attract better employees at lower cost, lower cost to introduce extensions) (Raggio and Leone, 2009).

Levels of Brand Value

Two levels of brand value are identified: *Current value* refers to the projected profits of the brand with its current ownership, management, strategies, capabilities, and resources (Raggio and Leone, 2009). *Appropriable value* refers to the theoretical level of brand value that could be achieved if all brand equity were fully leveraged through such activities as brand extensions, pricing, advertising, or distribution (Raggio and Leone, 2009). A firm can choose either to chase the appropriable value of its brand by continuing to own and invest in it; or the firm may choose to sell or license the brand to a firm with stronger brand management skills and a greater ability to leverage brand equity (Raggio and Leone, 2009). Estimates of current value set a floor on the price a firm would accept to sell its brand, while estimates of appropriable value set a ceiling on the price that an acquiring firm would pay.

Increasing Appropriable Value

Firms increase the appropriable value of their brands by building brand equity (e.g., through patents, endorsements, positive word-of-mouth, etc.), and then chase the higher appropriable value (thus increasing current value) of their brands by more fully leveraging existing brand equity or by selling or licensing the brand to another firm (Raggio and Leone, 2009). Thus, brand value varies depending on the owner (or potential owner) of a brand, and on the basis of the focal company's ability to successfully chase appropriable value, which is accomplished by leveraging existing brand equity.

Approaches to Brand Valuation

Three general approaches to brand valuation exist: cost-based, market-based, or future cash flow-based (Salinas and Ambler, 2009). Cost-based approaches rely on historical costs, but are hampered by the fact that not all brand "investments" are created equal. As such, ROI vary, and an estimate of money spent may not reflect the price to recreate the brand again in the future (Haxthausen, 2009). Market-based valuations attempt to compare a brand against other similar brands that have been sold, licensed, securitized, or valued for other purposes. But since management's ability to chase appropriable value varies, and because brands offer unique benefits to different target segments, market-based estimates must be adjusted for the idiosyncrasies of the focal brand and firm (e.g., size and growth rate of target segment(s), patents, trademarks, etc.). Finally, present values of future cash-flow procedures attempt to determine the portion of cash flows that are attributable to the brand as opposed to the underlying product or other factors (such as distribution efficiencies or other economies of scale/scope).

Uses of Brand Valuation

Brands are valued for a variety of reasons: brand valuation can be done to communicate the value of a firm's asset(s) to financial markets, potential buyers, or licensing partners, as a measure of the effects on consumer behavior (e.g., choice, willingness to pay a price premium, acceptance of brand extensions, etc.), or for management purposes (e.g., to gauge the effectiveness of brand managers and brand management strategies or tactics) (Haxthausen, 2009). Brand valuation procedures and estimates can vary depending

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

on the purpose of the valuation. For example, a cost-based approach may be informative if considering the decision to extend an existing brand versus develop a new one; market-based comparables may help justify licensing fees; and cash-flow based estimates may be important in the sale or securitization of a brand.

Bibliography

Barwise, P. (2003) Preface, in *Brands and Branding* (eds R. Clifton and J. Simmons), Bloomberg Press, Princeton, pp. xii–xv.

Haxthausem, O. (2009) Valuing brands and brand investments – key learnings and future expectations. *Journal of Brand Management*, **17**, 18–25.

Raggio, R.D. and Leone, R.P. (2009) Chasing brand value: fully leveraging brand equity to maximize brand value. *Journal of Brand Management*, **16**, 380–395.

Salinas, G. and Ambler, T. (2009) A taxonomy of brand valuation practice: methodologies and purposes. *Journal of Brand Management*, **17**, 39–61.

bundling

Shikhar Sarin

Definition

Bundling is defined as the practice of marketing two or more products and/or services in a single package (Yadav, 1994). As a strategy, bundling is becoming increasingly popular in marketing, not only for consumer products and services, but also for industrial products (Johnson, Herrman, and Bauer, 1999; Yadav, 1994; *see* PRICING STRATEGY; INTERNATIONAL PRICING OBJECTIVES AND STRATEGIES). However, despite its popularity, much remains to be learned about how to design optimal bundles (e.g., bundle size, bundle price, etc.). This discussion briefly highlights two elements critical to our comprehension of bundling: the process by which buyers evaluate bundles, and how buyers perceive the value of a bundle. We conclude with a brief discussion of some strategic implications of bundling.

Process of Bundle Evaluation

Early understanding of bundle evaluation was based on an additivity assumption (Yadav, 1994; *see* CONSUMER DECISION MAKING; CHOICE MODELS). That is, the overall utility of a bundle was assumed to be equal to the sum of the individual utilities of the bundled items. This assumption was subsequently discarded. Later research (e.g., Einhorn and Hogarth, 1985; Johnson and Puto, 1987) suggested that buyers evaluate a bundle by averaging the evaluations of individual items in the bundle, and that a weighted average model of bundle evaluation may be more appropriate (*see* CONSUMER DECISION MAKING; CHOICE MODELS).

Yadav (1994) notes that even in a small bundle of very few items, the amount of information to be processed by the buyer can be onerous. As such, buyers are likely to look for heuristics to simplify the evaluation task as a series of smaller (and simpler) evaluations (*see* CONSUMER DECISION MAKING; CHOICE MODELS). Yadav suggests that the anchoring and adjustment heuristic may enable the buyer to accomplish such an evaluation. Anchoring and adjustment involves an initial assessment (i.e., anchoring), followed by one or more adjustments. Using a series of experiments, Yadav (1994) showed that an anchoring and adjustment process of bundle evaluation involves three stages:

1. *Scanning stage*: where buyers determine which items are contained in a bundle. No evaluations are made at this stage.
2. *Anchor selection stage*: where buyers initiate the evaluation process by selecting one bundle item perceived to be the most important item for the evaluation task.
3. *Anchoring and adjustment stage*: where using the evaluation of the anchor as a starting point, the buyers initiate the evaluation of the bundle by evaluating the remaining items in the bundle in decreasing order of perceived importance. The buyers adjust their initial anchor evaluations upward or downward to reflect the new information.

Researchers (e.g., Einhorn and Hogarth, 1985; Johnson and Puto, 1987) note that bundle evaluations produced using the weighted average approach are very similar to those produced using the anchoring and adjustment method, suggesting that both these approaches capture the compensatory process of bundle evaluation in consistent but complementary ways (Yadav, 1994).

Valuation of a Bundle

In addition to the process of bundle evaluation, the perceptions of overall savings (*see* A FRAMEWORK FOR CREATING VALUE PROPOSITIONS; VALUE PROPOSITION) from the bundle are critical to the design of optimal bundles. Yadav and Monroe (1993) suggest that buyer's perceptions of overall bundle savings consist of two distinct perceptions of savings:

- perceived savings on the individual items if purchased separately, and
- perceived additional savings on the bundle.

Yadav and Monroe show that the additional savings offered directly on the bundle have a

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

greater influence on the buyer's perceptions of the transaction value, compared to the savings offered on individual bundle items. Others however suggest that more positive evaluations of a bundle are likely to result if the component prices are integrated into a single price, but the component discounts are segregated into a set of discounts (Johnson, Herrman, and Bauer, 1999). That is, buyers were likely to be more satisfied with bundles and more likely to purchase and recommend them, if the pricing information was bundled and the price discount information was unbundled.

Johnson, Herrman, and Bauer note that large gaps exist in our understanding of how best to present bundled price information to the buyer. Clearly, contradictory pricing guidelines such as those illustrated above, point to a need for additional exploration of this phenomenon.

Strategic Implications of Bundling

While much of the bundling research has focused on bundle evaluation and price bundling, recent research (e.g., Bakos and Brynjolfsson, 1999, 2000; Sarin, Sego, and Chanvarasuth, 2003) points to larger strategic implications of bundling. In addition to the increase in sales documented in earlier research, emerging research points to the use of bundling to market high-tech products. For example, Sarin, Sego, and Chanvarasuth (2003) suggest that bundling could be used as a strategy to reduce the perceived risk associated with the purchase of new-high tech products, and hence increasing their adoption and acceptance. They offer a series of recommendations for the optimal design of bundles for new high-tech products.

Similarly, Bakos and Brynjolfsson (1999) show bundling a large number of unrelated information goods to be very profitable. Bakos and Brynjolfsson demonstrate that the law of large numbers makes it much easier to predict the customer's valuation of a bundle of goods, as compared to their valuations of individual good when sold separately. This "predictive value of bundling" makes it possible for the seller to extract more value from each information good through greater sales, greater efficiency, and greater profits from a bundle of information goods, than can be attained from selling them separately (Bakos and Brynjolfsson, 1999, 2000). Bakos and Brynjolfsson suggest that large aggregators of information goods can be more profitable than small aggregators by offering a menu of different bundles, aimed at different market segments, increasing the impact of traditional price discrimination strategies.

Bibliography

Bakos, Y. and Brynjolfsson, E. (1999) Bundling information goods: pricing, profits, and efficiency. *Management Science*, **45** (12), 1613–1630.

Bakos, Y. and Brynjolfsson, E. (2000) Bundling and competition on the internet. *Marketing Science*, **19** (1), 63–82.

Einhorn, H.J. and Hogarth, R.M. (1985) Ambiguity and uncertainty in probabilistic inference. *Psychological Review*, **92**, 465–481.

Johnson, M.D., Herrman, A. and Bauer, H.H. (1999) The effect of price bundling on consumer evaluations by product offerings. *International Journal of Research in Marketing*, **16** (2), 129–142.

Johnson, M.D. and Puto, C.P. (1987) A review of consumer judgment and choice, in *Review of Marketing* (ed. M.J. Houston), American Marketing Association, Chicago, pp. 236–292.

Sarin, S., Sego, T., and Chanvarasuth, N. (2003) Strategic use of bundling for marketing new high-tech products: strategies for reducing consumers' risk perception. *The Journal of Marketing Theory and Practice*, **11** (3), 71–83.

Yadav, M. (1994) How buyers evaluate product bundles: a model of anchoring and adjustment. *Journal of Consumer Research*, **21** (2), 342–353.

Yadav, M. and Monroe, K.B. (1993) How buyers perceive savings in a bundle price: an examination of bundle's transaction value. *Journal of Marketing Research*, **30** (3), 350–358.

communications budgeting

Richard F. Beltramini

Communications budgeting refers to appropriating the human, financial, and physical resources needed to support the integrated marketing communications (IMC) activities of an organization, and then allocating these resources among the various IMC components (e.g., advertising, public relations, sales promotion, etc.). While this is a slightly complicated and inexact process, a number of commonly accepted budgeting methods are typically used today (see Arens, Weigold, and Arens, 2009 and others).

The "*percentage* (*of sales*)" method is perhaps the simplest communications budgeting method to employ. Organizations simply appropriate a percentage of their revenues to support the IMC function, ranging from a small percentage among, for example, industrial equipment or technology companies, to sometimes as much as 20% among alcoholic beverage or cosmetics manufacturers. Unfortunately with simplicity comes the slightly arbitrary assumption that all competitors are in identical positions, and typically the backward reliance on sales experience rather than driving future marketplace response.

Another communications budgeting method is the "*share* (*of market*)" technique in which organizations appropriate resources according to their desired share of market activity. Particularly for new-product introductions, a multiplier of desired share is often used initially (e.g., if a 10% share of market is desired, perhaps a 15% share of all resources expended by all competitive brands in its product category may be allocated for the product introduction period). Obviously, for this approach to communications budgeting to work, organizations must have knowledge of all competitors' IMC activities.

Finally, the "*objective*" method represents the most sophisticated communications budgeting approach, first developing quantifiable marketing communications objectives (what to accomplish), then strategies (what to do), and finally tactics (how to do it). For example, if an organization decides to "increase brand awareness by 10% among Midwest teens," it might want to strategically position the brand as offering the "best taste in the product category," and analyzing what media mix (e.g., broadcast, print, online, etc.) is needed to get that message communicated to its target audience. This approach builds up resources needed based on activities required to accomplish measurable market responses.

A number of factors obviously influence the selection of a communications budgeting method. For example, the reality of an organization's financial situation may ultimately limit the amount indicated by one of these methods, the seasonal timing and product life cycle of competing brands may dictate the need for accelerated (or decelerated) marketplace activity, and the availability of market intelligence may vary between simple intuition (or habit) and sophisticated empirical assessments of competitive activity and forecasted demand. In some cases, organizations might find it necessary to simply match what competitors are doing, while in other instances, a series of performance metrics might be used to map out a series of test-evaluate-retest refinements.

The most important element of communications budgeting, however, remains a solid understanding of the role of integrated marketing communications in effecting marketplace response. Communications activities alone or in combination do not simply *cause* market impact (awareness, interest, desire, or action). In fact, a myriad of controllable and uncontrollable factors complicate the relationship between say, a product advertisement run and the product sales afterward. While no one communications budgeting method is appropriate in all instances, individual selections require both a grasp of the theory behind how communications work and the practice of flexibility in the face of uncertainty.

Bibliography

Arens, W.F., Weigold, M.F., and Arens, A. (2009) *Contemporary Advertising*, McGraw-Hill/Irwin, New York.

Belch, G. and Belch, M. (2009) *Advertising and Promotion: An Integrated Marketing Communications Perspective*, McGraw-Hill/Irwin, New York.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Clow, K.E. and Baack, D. (2010) *Integrated Advertising, Promotion, and Marketing Communications*, Pearson Education, Inc., Upper Saddle River.

Shimp, T. (2010) *Advertising, Promotion, and other aspects of Integrated Marketing Communications*, South-Western Cenage Learning, Mason.

cannibalism

Dwight R. Merunka

Sales Cannibalism

New-product sales require market expansions, stolen sales from competing products, or cannibalism of sales of the firm's own products. Sales cannibalism is the extent to which a new-product's sales (or a current product's sales) come at the expense of other products offered by the same firm. It, thus, can be measured as the proportion of a product's sales drawn from product(s) of the same brand or made by the same manufacturer (Mason and Milne, 1994). In mature markets, new products do not attract new customers nor increase consumption rates among current customers. Therefore, a high proportion of sales of new products likely come from the cannibalization of the firm's existing products. In general, the extent to which cannibalization occurs depends on product substitutability or "substitutability in use," a concept that holds that products are regarded by consumers as interchangeable for some relevant purpose or usage situation.

According to this perspective, marketers generally consider sales cannibalism an outcome of improper product positioning and management (*see* POSITIONING ANALYSIS AND STRATEGIES). That is, cannibalism is something to avoid, because it is not desirable to lose market share or sales of profitable existing products when introducing a new one. Cannibalism represents a potential problem and a negative consequence of new-product development, as is well illustrated by channel cannibalization, or sales losses in existing channels as a result of penetrating a new one. In fact, fear of cannibalization has deterred many firms from deploying the Internet as a distribution channel (Deleersnyder *et al.*, 2002). However, in the context of radical product innovation, assuming that cannibalization is negative and should be avoided at all costs may be dangerous and risky.

Willingness to Cannibalize

Cannibalism can also be viewed as taking an initiative to supersede the firm's existing products before a competitor's entry attacks them (Conner, 1988). Therefore, the decision involves whether to cannibalize planned sales and profits on existing products by introducing a new product whose anticipated performances remain largely unknown. This conceptualization of cannibalism is particularly important for radical product innovations (*see* RADICAL INNOVATION). Because radical product innovations can result in substantial cannibalization of existing business, Chandy and Tellis (1998) argue that a positive disposition to cannibalize the firm's current resources represents a key variable to explain the firm's level of radical innovations.

Willingness to cannibalize is "the extent to which a firm is prepared to reduce the actual or potential value of its investment" (Chandy and Tellis, 1998, p. 475). It entails a multidimensional construct: Firms can cannibalize (i) investments in the form of assets (equipment, technology, knowledge, expertise), making prior investments obsolete; (ii) organizational capabilities (established procedures) that will render skills and routines obsolete; and (iii) sales, which require introducing new products that will diminish the sales of current products (Nijssen, Bas, and Vermeulen, 2005). A company's willingness to cannibalize explains why some companies develop more radically new products than do others (Chandy and Tellis, 1998). In this sense, it is a desirable trait that promotes radical product innovation and the long-term success of the firm.

Firms making substantial sales and profits with their current products may be reluctant to introduce new products if they fear they will be cannibalizing their current sales and prior investments. However, willingness to cannibalize drives radical product innovation, and firms must be willing to cannibalize before the obsolescence of their current products occurs – at which point there will be nothing left to cannibalize.

Bibliography

Chandy, R.K. and Tellis, G.J. (1998) Organizing for radical product innovation: the overlooked role of willingness to cannibalize. *Journal of Marketing Research*, **34**, 474–487.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Conner, K.R. (1988) Strategies for product cannibalism. *Strategic Management Journal*, **9**, 9–26.

Deleersnyder, B., Geyskens, I., Gielens, K., and Dekimpe, M.G. (2002) How cannibalistic is the internet channel? A study of the newspaper industry in the United Kingdom and The Netherlands. *International Journal of Research in Marketing*, **19**, 337–348.

Mason, C.H. and Milne, G.R. (1994) An approach for identifying cannibalization within product line extensions and multi-brand strategies. *Journal of Business Research*, **31**, 163–170.

Nijssen, E., Bas, H., and Vermeulen, P.A.M. (2005) Unraveling willingness to cannibalize: a closer look at the barrier to radical innovation. *Technovation*, **25**, 1400–1409.

competitive advantage: its sources and the search for value

Albert Caruana

The concept of competitive advantage highlights the close interlink between marketing and strategy. Strategy seeks to relate a company to its environment and to determine where the manager wants the business to go. To get there managers need to identify an activity or activities that can collectively offer value to customers. In the book *Competitive Advantage*, Michael E. Porter (1985, p. 3) makes this point clear and states that "Competitive advantage grows out of the *value* a firm is able to create for its buyers that exceeds the firm's cost of creating it."

Value lies at the core of marketing. This is what customers look for and are willing to pay for. The pivotal role of value has long been recognized in marketing and is reflected in the more recent definition of marketing, which the American Marketing Association (2007) declares as "the activity, set of institutions and processes for creating, communicating, delivering, and exchanging offerings that have *value* for customers, clients, partners, and society at large." Value is not an absolute but a relative concept recognizing that new competitive offerings may erode any value offering that the firm may have in the market (*see* CUSTOMER SOLUTIONS).

Value must necessarily be considered from the customers' perspective, yet unfortunately many organizations fail precisely in this regard. There are numerous reasons for this. A focus on manufacturing or distribution may lose sight of the ultimate beneficiary for whom these activities are being undertaken, while with service offerings value is often defined by the professional with little focus on customers. Indeed, medical practitioners look at their clients as patients, which by definition implies that they often need to be just that – patient!

This contribution highlights the salient role played by value and how it underpins the identification of viable sources of competitive advantage. It commences by taking into account the centrality of cost leadership and differentiation as two important sources of competitive advantage. These interact with particular target markets to provide three main generic strategies. Marketing advantage is then considered as a further basis of competitive advantage and while discussing cost leadership and differentiation advantage, a typology is investigated that explicates major sources that sustain each of these advantages. This is followed by consideration of how the value chain enables the identification of activities that underline a particular value offering that underpins the different types of competitive advantage that the firm can provide. Finally, the main generic strategies are discussed while attention is also drawn to other perspectives on competitive advantage. In conclusion, the point is made that the various available frameworks are best viewed as aids and not as substitutes for management thinking on competitive advantage.

Generic Strategies as Sources of Competitive Advantage

To provide value once is not sufficient and competitors will act to erode any advantage a firm may have, often via imitation and *innovation* (*see* INNOVATION DIFFUSION). For a competitive advantage to be real it must endure over time. When an advantage can be maintained over time it is said to be *sustainable* and managers can talk of a sustainable competitive advantage. For sustainability to exist, the relative advantage that the positioning achieves needs to have meaning for target customers (*see* MARKET SEGMENTATION AND TARGETING).

The challenge for managers is to identify the sustainable competitive strategy that will underpin the direction that they will pursue for their firm or for their product. Porter (1985) argues that there are two main sources of competitive advantage that he terms *cost leadership* and *differentiation*. Competitive advantage is possible by either pursuing a low-cost positioning or by underscoring differentiation that involves emphasizing some perceived uniqueness that exists or can be fostered among customers. Alternative positioning for the firm can be achieved relative to other players in an industry depending on the choice of strategic target (or competitive scope) in terms of whether the firm pursues a broad industry-wide target or whether it targets a narrower segment.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Competitive scope		Competitive advantage: Lower cost	Competitive advantage: Differentiation
	Broad target	Cost leadership	Differentiation
	Narrow target	Cost focus	Differentiation focus

Figure 1 The three generic strategies. Source: Porter (1985, 12).

These two dimensions of competitive scope and competitive advantage are utilized in a matrix to highlight the three potential successful generic strategic approaches achievable. Depending on the target market focused on, it is possible to identify three successful generic strategies that can enable the firm to outperform competing firms in an industry. These include the following:

1. Overall cost leadership
2. Differentiation
3. Focus.

Overall cost leadership and differentiation are positioning decisions that can have meaning for broad targets, while narrow targets require focus on either cost or on differentiation as depicted in Figure 1. Each positioning can provide a competitive advantage and represents a unique value offering to customers that allows different firms to compete in the same industry without being in direct competition with each other at most times.

It is clear that no firm, in whatever industry, can offer value and be all things to all people. To achieve a sustainable competitive advantage, the firm needs to make choices that can offer value to customers and that value must not be easy to imitate by competitors. Such value may arise from offering customers great service as a result of customer-centric delivery processes that are not easy to replicate. Value may also come from very efficient operating systems that are able to squeeze out costs and pass some of that value to customers. The search for value offerings is ongoing in all competitive industries.

The airline industry offers a good example of these strategies at work. In this industry, there has been an ongoing struggle to target and differentiate broad groupings. This has been pursued by players such as British Airways who have used full service differentiation throughout the organization in their Putting People First programs, calling themselves "the world's favorite airline" in an attempt to differentiate on a basis that goes beyond the British nationality factor that can offer only limited differentiation. On the other side of the cost leadership matrix, the entry of low-cost airlines such as Ryan Air and Easyjet underlines the pursuit of a positioning that has witnessed the erosion of customers away from the many airlines that previously sought a differentiated positioning. In addition, within the industry, there are many examples of small regional airlines that necessarily target narrowly and can compete on cost with the more differentiated airlines on particular routes. Before its demise, Volareweb that operated from Italy to select European routes sought to compete on price. However, it is also possible to target narrowly by focusing on routes that are not well served as between Rome and Lugano, Switzerland where Darwin Airlines can pursue a more service-quality-differentiated focus strategy.

Best (2004) recognizes cost advantage and differentiation advantage as the two primary sources of competitive advantage but also adds marketing advantage as a third source of competitive advantage. This involves marketing activities that make it possible for the firm to dominate competition on the basis of sales, distribution, or brand recognition, or some combination of these three. In addition, these three categories are used to provide a useful typology that clarifies the activities by which a business can achieve a competitive advantage. The typology seeks to drill down on each of the sources of competitive advantage by considering three variations under each category (Figure 2).

Cost advantage. The competitive advantage typology is useful and its three main variations are now discussed. We start by first looking at cost advantage, which arises when the business is able to generate economic value by having

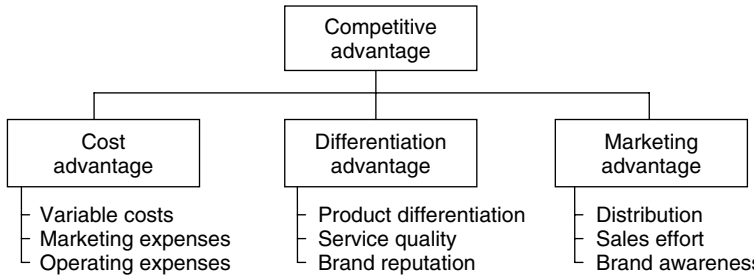


Figure 2 Major sources of competitive advantage. Source: Best (2004, 140).

costs that are lower than those of competitors. The three main sources of cost advantage are reviewed.

Lower variable costs. Lower variable costs can result in lower prices or higher margins relative to competitors. One of the most common ways in which a firm can achieve lower variable costs is via economies of scale. A firm with a large MARKET SHARE is able to achieve purchasing efficiencies and lower the cost of production and can enable costs to be split across a larger output (*see* BRAND GROWTH STRATEGY). Such a position precludes competitors from gaining market share as it will be difficult for such firms to achieve the necessary volume to match unit costs. Large volume production can also benefit from a learning curve effect where not only do unit costs decrease because of growth in the volume of production but there is also a further drop that results from a learning effect. This comes about because, with practice, people and organizations become better at their tasks, thereby increasing the efficiency of production making it possible to cut unit costs even further. Typically, learning curves follow a negative exponential distribution.

Marketing cost advantage. Marketing cost advantage results from a marketing cost scope effect whereby the various costs of marketing can be divided over a broader range of outputs as a result of extensions in product breadth and/or width. The pursuit of umbrella branding is a case in point. In contrast to individual product branding, umbrella branding is especially widespread among service firms, including banks. Such firms can gain from economies of scope, in this case, with respect to advertising, whereby a single advertising campaign can be employed to promote an entire range of product offerings. It is possible to derive a marketing cost advantage not just from lower advertising costs but also from other marketing activities.

Operating cost advantage. Operating costs refer to the expenses involved in running a business that is not directly related to the production and sale of the current line of goods or services. These include such costs as the general and administrative expenses involved in running a business which, if they can be lowered, can also provide a low-cost advantage. The new Airbus A350XWB scheduled to come into service in 2013 is a wide-bodied aircraft that competes with the Boeing 777 and 787. The manufacturer claims that the A350XWB will be able to offer airlines buying the aircraft "a 20 per cent lower cash operating costs per seat than competing aircraft in this size category and fuel efficiency improvements of up to 25 per cent per seat." In addition, it has been reported that Airbus has been looking at putting a 10-abreast instead of a 9-abreast seat configuration allowing the operating cost per seat to be such that it could not be matched by any other aircraft.

Since not all businesses are able to have some type of cost advantage, managers can consider seeking a differentiation advantage as their source of competitive advantage. *Differentiation* (*see* POINT OF DIFFERENCE AND PRODUCT DIFFERENTIATION) represents a second alternative competitive advantage to the firm that arises when the business is able to generate economic value by giving a product offering that customers prefer over competitors' product offerings. A differentiation advantage is possible

on practically any basis as long as it can meet some customer need or needs better than those of competitors. The three main sources of differentiation advantage are elaborated.

Differentiation advantage.

Product advantage. The product offering provides various attributes that allow for differentiation. While an exhaustive list is necessarily elusive, it is possible to provide a few examples that can give an indication of the possible alternatives available. Thus appearance, design, and features have been used by Apple to strongly differentiate their offerings. Product features have been used effectively by Toilet Duck where its original duck-shaped dispenser for toilet cleaner designed to reach under the rim of toilets was able to provide a distinctive *positioning* (*see* POSITIONING ANALYSIS AND STRATEGIES) in the detergent market. Product benefits are also widely used for the purpose of differentiation with the toothpaste market offering a good example with brands like Sensodyne for sensitive teeth and Crest against tooth decay clearly emphasizing different benefits to consumers.

Service advantage. Attribute characteristics that can provide a differential advantage to tangible products can also be leveraged for intangible or *service* products (*see* CONSUMER BEHAVIOR AND SERVICES MARKETING). Indeed most offerings in the market can be seen as representing a combination of the tangible and the intangible. However, the aspect of service quality is an area that has enabled differentiation among the more intangible offerings in the market and has received much research attention. Often defined as the difference between expected and actual performance, it has been shown that service quality has beneficial consequences in terms of customer satisfaction, loyalty, and ultimately profitability. Service quality is seen as encompassing dimensions of reliability, responsiveness, assurance, empathy, and tangibles. Service quality is demanding to be put in place and, if done successfully, can provide a service firm with a differential advantage. Singapore Airways is the only one in a handful of airlines that gets a five-star rating from SKYTRAX, the airline and airport quality certification firm, which confirms that Singapore Airways provides a high-quality service experience to passengers.

Reputation advantage. Reputation can be considered at both a product and at a corporate level. It is an aspect that is also able to provide the firm with a competitive advantage. Like service quality, building a reputation is demanding and it cannot be easily copied. While brands can be thought of as symbols around which relationships can be built between sellers and customers, brand and corporate reputation is essentially a feeling or a disposition to respond favorably or unfavorably to a brand or company. However unlike brands, reputation extends beyond customers and can vary by different communities or publics. Thus, while Abercrombie & Fitch is undoubtedly a strong brand that sells a lifestyle image popular with teens and young adults, the brand and its corporate reputation among some noncustomer groups and other communities have seen some of these noncustomer publics raise questions about such things as the quality of the clothes and the advertising approach adopted by the brand. On the other hand, Google is able to have a strong reputation among different publics. In its reputation, it possesses an asset that can provide various benefits to the firm. Therefore, its reputation advantage can result in increased sales; can foster more credible advertisements; can improve perceived product quality; can produce higher customer loyalty; can add value; attract higher margins, and increase cash flow and profits. It can also attract investors and high-quality job applicants.

Marketing advantage. Besides pursuing a cost or a differentiation advantage as their source of competitive advantage, it is possible for a firm to gain competitive advantage by leveraging a combination or a particular marketing strength relating to sales coverage, distribution, or brand recognition. These three aspects are further elaborated as follows:

Channel advantage. The ability to gain distribution coverage or to deny coverage to competitors is an important competitive advantage that goes beyond cost or a differential advantage. Cisk, a dominant market share lager

brand in Malta that came under increased competition from a new local brewery that bottled Lowenbrau, was able to counter the new competition by effectively denying access to the distribution network to the competitor. It succeeded by using pricing and sales promotion activities directed at the distribution.

Sales force advantage. There are four key strategic sales force challenges that need to be met, which, if they are successfully tackled, can provide the firm with a competitive advantage. These relate to: how many sales people should be employed; whom they should visit, and with what frequency, and how the sales force should be organized. If met properly together, these can provide the firm with a sales force that is effective and can also afford superior coverage to that of competitors with sales forces that are larger or of comparable size.

Brand awareness. Brand awareness refers to the level of recognition that customers have of a brand together with knowledge of the specific product category to which it appertains. With many fast-moving consumer goods where the difference among products is marginal, brand awareness plays a very important role. Colas are a case in point. The differences in the colas at a taste level are relatively minor but brand awareness for the likes of Coca-Cola and Pepsi, together with the image they portray, results in higher sales for these brands and provides a significant barrier that precludes other colas from gaining market share. The brand awareness of these two brands effectively provides them with a competitive advantage irrespective of possible superior qualities in terms of attributes such as taste or price that other competing colas may have.

Value Chain, Value Provision and Competitive Advantage

Porter (1985) sets out the concept of the value chain as a useful framework that describes how the organization undertakes certain activities that are able to create total value and uphold a particular competitive advantage. Any firm can be thought of as a set of limited resources that makes use of activities to provide value that enables the firm to earn a margin. Margin is generated when created value exceeds costs.

The value chain allows management to systematically consider the various activities involved. The primary activities in the value chain that enable the firm to make its offering are inbound logistics, operations, outbound logistics, marketing and sales, and service, while support activities consist of firm infrastructure, human resource management, technology development, and procurement. The latter three can support each of the primary activities while firm infrastructure acts to support the entire chain (Figure 3).

The value chain highlights the fact that value can be provided not just from the products and brands (*see* BRAND VALUE) that the firm makes available but also from effective procurement and the channels of supply and distribution. The activities that are brought together in the value chain allow value to be provided to customers, determining the differentiation that can take place, and the costs that will ultimately result. One needs to understand how these activities are being utilized, how they are controlled, and how they interact, with a view to devising ways so that they can be used to gain competitive advantage.

Analysis of the elements of the value chain underlines the search for synergetic effects that can result from the extra benefit that can be derived by the firm by not only looking at the elements of the value chain but also by bringing together the added value the different linkages in the value chain are able to provide. Competitive advantage results because the firm is able to undertake these activities more effectively or at a lower cost than competitors.

The value chain is not without its critics and it has been argued that the mental framework behind it appears to assume a manufacturing rather than a service entity. In addition, it appears to betray an assembly-line likeness in its linear, sequential, and unidirectional focus.

Alternative Generic Strategies

Porter (1980, 1985) cautions against a failure to opt for one of the three alternative generic strategies and describes firms that fail to pursue one of the three generic strategies as "stuck in the middle." It has been suggested that such

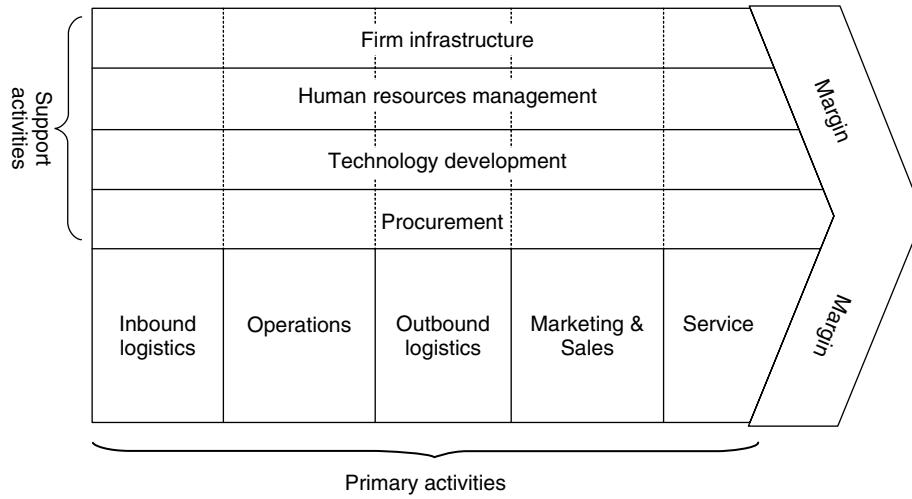


Figure 3 The generic value chain. Source: Porter (1985, 37).

firms are unable to use their resources effectively and consequently can be expected to exhibit poor profitability. While accepting temporary situations where the pursuit of both a cost leadership and a differentiation strategy can be possible, firms are encouraged to look at their resources and determine the strategic orientation that can best be matched to their capabilities. It is therefore argued that firms cannot simultaneously pursue a cost leadership and a product differentiation strategy. Rolex, which pursues a narrow differentiated focus positioning, is a case in point. Its structures, management control, and compensation policies preclude the pursuit of a cost-focused positioning.

However, it has been suggested that the generic strategy positions may not always be as mutually exclusive as suggested. Hill (1988, 401) has argued that "the circumstances in which the simultaneous pursuit of differentiation and low cost make sense are common, depend on more factors than previously highlighted, and may well lead to the establishment of a sustainable competitive advantage." A contingency framework is therefore proposed where, in the circumstances of certain emerging and mature industries with particular characteristics, managers can successfully pursue a sustainable competitive advantage that can frequently involve the simultaneous pursuit of differentiation and low-cost strategies.

A similar point is made by Miller (1992) who argues against the notion of "stuck in the middle" and holds that a strategy that straddles both positions is viable. Therefore, Caterpillar Inc., which pursued a differentiated strategy that focused on providing earth-moving equipment with distinctive characteristics in terms of high precision and durability while paying less attention to efficiency and economy, enabled Japanese firms to undercut them on price by 30% – the point being that overspecialization and differentiation of the offering can omit important attributes that customers are also looking for and that may afford an opportunity that can be exploited by competitors. Toyota would seem to be an example of a firm that has been able to pursue both a cost leadership and differentiation strategy simultaneously. This is possible because some bases of differentiation have been able to also lower costs. The structures, policies, and controls employed are complementary, and enable such a stance to be possible.

The debate as to the relevance of generic strategies is ongoing. A number of authors have suggested that cost leadership and differentiation strategies are too generic and are limiting. Kim and Mauborgne (2005) look at what they call *value innovation* and argue that a firm should look outside their current paradigms to find new value propositions (*see* A FRAMEWORK FOR

CREATING VALUE PROPOSITIONS) in new market space or "Blue Ocean." The pursuit of a Blue Ocean strategy argues against a firm competing in the current market but rather to identify a market demand that has not yet been met.

More recently, Gurău (2007) notes the increased complexity of the current competitive environment that firms operate in and argues for the need for stronger relationship marketing among firms that can build a competitive advantage based on customer loyalty. He envisages management seeking to move their relationship with customers along a continuum from a weak to a strong relationship and suggests employing different aspects of generic strategies, starting with cost leadership in situations of weak relationships and progressing to product quality-based differentiation, support services-based differentiation, a niche strategy, personalized marketing, and ultimately value cocreation in the strong relationship condition on the continuum. Gurău sees relationships, which enable individual customers to cocreate unique experiences as a result of high quality interaction, as increasingly becoming an important source of competitive advantage to the firm.

While success can indeed come about by the adoption of one of the generic strategies, there are circumstances where a mixed strategy that combines elements of both cost leadership and differentiation strategies or a progression of generic strategies on a relationship continuum, can possibly represent a viable alternative way forward. Porter's generic strategies are not a substitute for management thinking. Matrices are by their very nature simplistic and reality is often not two-dimensional but multidimensional and complex. Rather, generic strategies and their related matrices should be viewed as aids to thinking that can provide useful insights and indications of possible options for management action in the pursuit of competitive advantage.

Bibliography

American Marketing Association (2007) http://www.marketingpower.com/Community/ARC/Pages/Additional/Definition/default.aspx (accessed 24 October 2009).

Best, R.J. (2004) *Market-Based Management, Strategies for Growing Customer Value and Profitability*, 3rd edn, Prentice-Hall Inc, New Jersey.

Gurău, C. (2007) Porter's generic strategies: a re-interpretation from a relationship marketing perspective. *The Marketing Review*, **7** (4), 369–383.

Hill, C.W.L. (1988) Differentiation versus low cost or differentiation and low cost: a contingency framework. *Academy of Management Review*, **13** (3), 401–412.

Kim, W.C. and Mauborgne, R. (2005) *Blue Ocean Strategy: How to Create Uncontested Market Space and Make the Competition Irrelevant*, Harvard Business School Press, Massachusetts.

Miller, D. (1992) The generic trap. *The Journal of Business Strategy*, **13** (1), 37–41.

Porter, M. (1980) *Competitive Strategy: Techniques for Analyzing Industries and Competitors*, The Free Press, New York.

Porter, M. (1985) *Competitive Advantage: Creating and Sustaining Superior Performance*, The Free Press, New York.

competitive analysis

Venkatesh Shankar

Introduction

Competitive analysis refers to the evaluation of strategies relating to a focal brand, product, strategic business unit (SBU), or firm relative to the current and potential competitors in its market or industry. Competitive analysis can be done at the product-market level, the SBU level, or the firm level. The main purpose of competitive analysis is to develop business and marketing strategies that take into account competitive considerations.

Competitive analysis is typically performed on an ongoing basis. Competitive analysis is conducted before the formulation and development of positioning or value proposition strategy and marketing mix strategies. It is also undertaken for a new brand or product, in particular, prior to its launch in a new market. As such, it is the basis for the development of competitive marketing strategy.

The terms, *competitive analysis* and *competitor analysis*, are used interchangeably. Strictly speaking, competitive analysis is related to but different from competitor analysis. Competitor analysis involves the study of competitors' strengths and weaknesses. It provides key inputs to competitive analysis. Competitor analysis focuses on competitors, whereas competitive analysis focuses on own brand, product, SBU, or firm.

In formulating brand or product positioning strategy, customer perceptions of the focal brand's strengths and weaknesses relative to competitors in a given product market form the inputs for deriving a perceptual map for that market. Using this information and data on the location and size of customer preferences, a brand or product can develop its positioning strategy. Thus, competitive analysis helps a brand or product differentiate itself from competing brands or products.

A good competitive analysis should evaluate the focal brand, product, SBU, or firm relative to both existing and potential competitors. Future competitors include marketers of substitute products and those with alternative technologies capable of satisfying customer needs in the relevant markets or industries. Competitive analysis is performed at intervals of a year or less.

At the industry level, competitive analysis is typically performed using the five-forces model (Porter 1980, 1985). The five forces that shape industry profitability include threat of new entrants, threat of substitute products, bargaining power of suppliers, bargaining power of buyers, and competitive rivalry. In this analysis, the factors that influence each force are identified and an assessment is made on the level of influence of each force on industry profitability. An overall evaluation of the attractiveness of the industry is reached on the basis of the relative influences of the five forces.

Competitive analysis is also a useful tool for deciding whether, when, and how to introduce a new product in a market. Firms typically decide to launch products when their strengths relative to their competitors outweigh their relative weaknesses in markets where opportunities surmount threats.

An Example of Application of Competitive Analysis for a New Product

Consider the example of competitive analysis for iPhone's marketing strategy before its introduction in 2007. iPhone's key strengths can be analyzed along six dimensions: innovativeness, compatibility, ease of use, brand awareness, price, and quality. The iPhone had an innovative touch screen that is patented and unmatched by other mobile products at the time of introduction. It offered multiple functions in a single mobile device. It was compatible with iTunes, other Mac products and operating system software tools, and other Apple products like Apple TV. It allowed wireless connectivity to the big screen. Its touch-screen interface was easy to use and its operations were intuitive. It was radically different from other phones or personal digital assistants (PDAs) in that it recognized multi-finger gestures common with a normal human hand. Apple was well known for cool essential gadgets like iPods and Macintosh. iPhone initially planned to sell at a reasonable price for its value. The first iPhone was of high

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

quality as it had one of the brightest and most scratch-resistant screens.

iPhone's key weaknesses were related to targeting, price, and user interface. The iPhone and Apple products, in general, had not been targeted at business people, whom most smart phones had targeted. It did not have a reputation as being compatible with the corporate world. iPhone did not offer lower priced models for price-sensitive consumers. Finally, the touch-screen interface suffered from the problem of "gorilla arm," a condition in which the long-term use of a flat, solid surface for input is uncomfortable for users.

iPhone's key competitors at the time of its launch included Blackberry, Palm, and Microsoft. Although Blackberry was targeted at corporate users and iPhone was directed at individual users, in terms of functionality and features, Blackberry was a close competitor. Blackberry's major strengths were that it catered to a large pool of corporate users and offered reliable email and voice capabilities on the go. The mobile hi-speed Internet connectivity was fairly robust and it was the first to offer corporate email on the mobile device that boosted office productivity. However, it had its own drawbacks as well. The scrolling feature and the keypad for data entry were cumbersome with limited capabilities.

Palm was also viewed as a direct threat in the handset market. With a long history in PDA market and experience in developing software for mobile devices, Palm is a well-known brand for businesspeople. Its software was well established and compatible with many products in that market. There were significant costs for Palm users to switch to iPhone. Palm's major weaknesses were that it was not known for ground-breaking innovations or consumer products. The Palm Treo had a lower quality perception and its interface harder to use, being from an older generation of products and unfamiliar to Apple customers. Furthermore, according to product reviews, Treo's keyboard was perceived to be small and difficult to use, especially for accessing the Internet. Although some users may have preferred having a keyboard built into the device, potential users showed a preference toward using the touch screen.

Microsoft was considered an indirect threat because it provided only the operating system for mobile devices and was in the process of introducing multifunctional mobile devices. Microsoft's major strengths were its financial flexibility and high cash holdings that could be used to develop new products and react aggressively to competition. Microsoft software, the Windows CE operating system for mobile devices, was already "proven" in the marketplace. Microsoft had global brand recognition in software products, in particular, in business markets. Businesspeople were likely to order its products on the strength of its name, whereas Apple did not make devices for business use. Microsoft's key weaknesses were its failed portable music player, Zune, the lack of cool image typically associated with Apple's products, and unexciting customer satisfaction rates.

On the basis of an analysis of iPhone's strengths and weaknesses relative to those of its competitors, Apple decided to launch iPhone by targeting individual users, positioning around its touch screen, easy-to-use innovative features, and compatibility with multiple software applications, projecting a hip image, pricing it at $599 at the retail, and selling it exclusively through AT&T, the largest wireless operator and well-entrenched brand name. The competitive analysis enabled iPhone capture the hearts of the market but it also revealed problems with its pricing. Within months of the launch, Apple had to drop iPhone's price by about a third and even offer credit notes and refunds to the initial buyers.

Game Theoretic Competitive Analysis

Advances in game theory have provided new tools for competitive analysis. The concepts of game theory can be applied to study competition in marketing variables and predict market outcomes. In game theoretic models, players (firms and/or consumers) play a noncooperative game in which they choose their strategies. A firm's strategy is a complete specification of its actions in all contingencies.

Game theoretic models assume that firms are rational and intelligent. Rationality refers to the firm's maximization of subjective expected utility (profits) and intelligence refers to the

firm's recognition that its competitors are rational. Game theoretic models can be analyzed by solving for equilibrium strategies. Equilibrium is that condition from which no competitor has an incentive to unilaterally deviate from its strategy. In other words, equilibrium strategies are those for which the competitors obtain the best outcomes in the given conditions.

The first step in game theoretic model is specification of the rules of the game. These rules include the number of competitors, the set of feasible strategies or decision variable(s), the profit function, whether it is a one-period (one shot) or multiperiod game, whether one competitor is a leader or follower in the decision variable(s), and the information known to the competitors at different stages of the game. Some information such as which competitor moves first is known to all the competitors and is known as *common knowledge*. If all information is common knowledge, then the game is known as a game with *complete information*. Some information such as manufacturing cost can be known to the focal firm but not to its competitors and is therefore, labeled *private information*. In this case, some rules of the game are not common knowledge, so this game is known as a game with *incomplete information*.

Information can be symmetric or asymmetric. In the case of information symmetry, all competitors are similar in their information knowledge. In the case of information asymmetry, Competitor A may know information on a variable (say cost) about Competitor B, but Competitor B may not have the information on the same variable about Competitor A.

Different types of game theoretic models exist for analyzing different conditions. For example, a game in which one competitor is a leader and another is a follower is known as a *Stackelberg game* and the corresponding equilibrium is known as the *Stackelberg equilibrium*. In contrast, the equilibrium corresponding to a game in which all competitors move simultaneously is known as the *Nash equilibrium*. For the possible existence of both these equilibria, see Shankar (1997).

Game theoretic models exist for analyzing competition in product, price, distribution channels, advertising, and promotion (for a detailed review, see Moorthy 1985). While much of game theoretic analysis is based on noncooperative interdependence among competitors, there is also cooperative game theory that examines the collusive behavior of firms in certain situations. The predictions of game theory models can be tested through empirical analysis. While game theoretic models are powerful tools for competitive analysis, they have some limitations as well. These limitations include unrealistic assumptions about the game and inability to capture real-world phenomenon through tractable models.

Competitive Analysis and Reaction to New-Product Entry

A new product's market success depends on its entry strategy as well as the response of competitors to its entry. Competitive analysis relating to reaction to new-product entry involves a solid understanding of the determinants of new-product entry strategy, its interrelationship with competitor response, and the drivers of competitor response – all of which can be analyzed through a framework proposed by Shankar (1999). According to this framework, a new product's entry strategy and competitor response are determined by entrant, incumbent, and market/industry characteristics. A new-product entry strategy is reflected by decisions on marketing variables such as product line length, price, advertising, and distribution. New-product entry strategy involves competitive considerations, and competitive advantages for pioneers and late movers under different conditions have been well documented (e.g., Shankar, Carpenter, and Krishnamurthi 1998, 1999). The literature on order of market entry in general is extensive (for reviews of order of entry effects in physical and electronic markets, see Kerin, Varadarajan, and Peterson, 1993; Varadarajan, Yadav, and Shankar, 2008).

To better analyze competitive response to new entries, we can use normative and descriptive models of response. A number of normative or game theoretic models of incumbent reaction exist. A seminal analytic model by Hauser and Shugan (1983) shows that it is optimal for an incumbent to decrease advertising and distribution spending in response to a new-product entry

in a static market. Their model also shows that if the consumers' tastes are uniformly distributed, then the optimal competitive strategy is to decrease prices, improve product quality, and reposition the brand more along the incumbent's strength.

A decoupled response function (the effects of advertising and distribution are decoupled) model by Kumar and Sudarshan (1988) shows that the optimal defensive strategy for incumbent competitors is to decrease price, advertising, and distribution when tastes are uniformly distributed and the market is static. A refinement of this model by Gruca, Kumar, and Sudarshan (1992) shows that the incumbent competitor's optimal response depends on its market dominance, with dominant brands (those with a market share of 50% or more) needing to increase spending and nondominant brands reducing spending.

Descriptive models of competitor response to new-product entry are typically based on econometric analyses of data from different markets. A simultaneous equation model by Robinson (1988) examines the effects of entry strategy, incumbent characteristics, and industry characteristics on incumbent marketing reactions in oligopolistic markets. Analysis using data from the Strategic Planning Institute on 115 new-product entries shows that incumbents do not react in the first year after entry and about half of them are passive in the second year after entry. The analysis, however, reveals that reactions in the second year after entry are more aggressive.

In another descriptive model, Gatignon, Anderson, and Helsen (1989) examine how established competitors react to a new-product entry in an over-the-counter gynecological product and in the airline industry. Their model focuses on the effectiveness of a current competitor's marketing mix instruments. The results show that incumbent firms retaliate with their more powerful marketing weapons and accommodate with less marketing instruments.

Using a game theoretic model that is both descriptive and normative, Shankar (1997) explains the pioneer's reactions and predicts its shift in marketing mix allocation upon new entry in a market. His competitive analysis involves study of equilibrium competitor reactions under Nash and different leader–follower games and an empirical illustration of analytical results with empirical analysis of pharmaceutical data. The analysis shows that the type of competitive game and the anticipated impact of the late mover on the pioneer's margin and elasticities are two critical factors that significantly affect the pioneer's decisions, in addition to the pioneer's characteristics and the market conditions. The findings reveal that a follower (leader) role in a marketing mix variable, a static (growing) market, a decrease (increase) in own elasticity and margin, generally lead to accommodation (retaliation) in that variable. This model underlines the need to examine a combination of factors such as competitive structure, the impact of the late mover on the pioneer's elasticities and margins, competitors' marketing mix elasticities, and own elasticities to understand shift in competitors' marketing mix allocation.

Firms can also perform an analysis of the determinants of reaction time of competitors to a new-product introduction. A model by Bowman and Gatignon (1995) estimated on the profit impact of market strategies (PIMS) data identify significant predictors of response time. The findings show that response time is positively related to customer switching costs, the entering firm's market share, and new-product development time, but negatively related to market growth, the reacting firm's market share, and the rate of technological change. Another study of response times to easily imitated new products by MacMillan, McCaffery, and van Wijk (1985) shows that response time is strongly related to visibility, radicality, complexity, perceived potential, strategic attack, and organization misfit.

Firms can study the effectiveness of different response times to new-product entries. Gatignon, Roberts, and Fein (1997) find that faster reactions generally have a positive impact on the perceived success of the defending firm but too broad a reaction (in many marketing mix instruments) is less effective. Importantly, the results show that the ability of an incumbent to maintain its market position depends on industry characteristics and the magnitude of competitive threat from the new product.

Further insights into the magnitude and timing of incumbents' responses to competitive new-product entries are available. A delayed response may be an optimal or efficient strategy when the new product's quality is uncertain because an immediate reaction in the form of a lower price by the incumbent firm may cause consumers to believe that the new product's quality is high (Kalra, Rajiv, and Srinivasan, 1998). A good competitive analysis in this regard should treat customers as strategic agents.

Other factors that influence the speed and breadth of retaliation include rival product's innovativeness, market growth, market concentration, and competitor size. Kuester, Homburg, and Robertson (1999) show that innovativeness and firm size slow incumbent's reaction time, market growth expedites retaliation in the product mix, and concentrated markets induce firms to react less strongly on the product mix and exhibit slower reactions.

Analysis of competitor reactions in industrial markets reveals new insights into business-to-business competitive analysis. In an empirical study of 509 new industrial products launched in the United States, the United Kingdom, and the Netherlands, Debruyne *et al.* (2002) find that two-thirds of new product launches meet reactions by competitors after their launch, primarily in the form of price changes. The analysis further reveals that product assortment and promotional changes are less frequent; distribution policy modifications occur very rarely; competitors fail to respond to radical innovations and to new products targeted at niche markets; competitors react if a product can be assessed within an existing product category; and competitors react to new products with high communication spending and in high growth markets. Thus, competitive analysis in industrial markets should consider the characteristics of the new-product launch strategy and whether the new products are radical or incremental innovations.

Competitive analysis of responses should also consider signaling characteristics of the new entries. Three market signals, hostility, consequence, and commitment are important to consider (Heil and Walters, 1993). Research suggests that, in general, the hostility signal has the greatest impact on reaction strength followed by the consequences signal and the commitment signal.

What are the factors that determine whether an incumbent will react to a new product announcement signal from a competitor, what are the variables that affect the aggressiveness of the incumbent's reaction, and what are the factors that influence whether the reaction will be in product or other marketing instruments? Robertson, Eliashberg, and Rymon (1995) address these questions based on survey data from the United States and the United Kingdom. Their results show that signal hostility is positively related to the occurrence of a reaction; an incumbent is more likely to respond when its fixed commitments in the product category is high and when the industry has high patent protection; and aggressive reactions are more likely under conditions of high signal credibility and in industries with high patent protection. Their findings further show that, in industries where patents are relevant, firms are less likely to make product changes and introductions; companies with high fixed commitments in the product category tend to react with other marketing mix elements, instruments, and hostile signals do not induce changes to product mix.

Competitors react to perceived market signals from a new product's launch decisions (broad targeting, penetration pricing, advertising intensity, and product advantage). Analysis by Hultink and Langerak (2002) on the effects of three perceived market signals on the strength and speed of competitive reaction shows that incumbents consider high advantage new products to be hostile and consequential and regard penetration pricing and intense advertising to be hostile, especially in fast-growing markets. Their results also reveal that broad targeting is not perceived to be hostile, especially not when used by entrants with an aggressive reputation and that perceived signals of hostility and commitment positively impact the strength of reaction while perceived consequence signal positively impacts the speed of reaction.

How do we analyze competitive effects of new-product introductions and deletions and proactive and reactive actions by competitor firms? Shankar (2006) analyzes actions and actions in product line, price, and distribution,

which include both anticipation and reaction components of market leaders and followers. His results show that market leaders have higher product line elasticities than followers, and that unlike followers, leaders have greater reaction elasticities than anticipation elasticities. These findings are useful in predicting which competitors will react in what marketing mix instrument.

Competitive analysis involving new-product entries and competitor responses have been based primarily on their impact on sales and market share. However, it is important to analyze their effects on the profits and firm values of different competitors, including the introducing firm. What are the competitive entry strategies that ensure both sales and profit growth and firm value? What are the competitor response strategies that result in market growth, but improved market position and profitability for incumbents? We do not yet have good answers to these questions.

Other Issues in Competitive Analysis

Competitive analysis can be enriched by including three topics that are becoming increasingly important for firms in a growing competitive landscape. These topics are competitive benchmarking, convergence analysis, and coopetition.

Competitive analysis can be enhanced by including competitive benchmarking. Competitive benchmarking involves analysis of the focal brand, SBU, or firm relative to the strongest competitor(s) and is typically done for strategies such as product strategy, promotion strategy, marketing communication strategy, sales force strategy, distribution strategy, and pricing strategy. Competitive benchmarking is closely related to analysis of best practices.

Competitive analysis is complex in converging markets. Convergence is a process by which the boundaries across industries, businesses, markets, geographies, or customer experiences become blurred, creating new competitors for firms (Shankar, Ancarani, and Costabile, 2010). Convergence competition increases in an economic downturn, involving new technologies, modified customer needs, and unconventional competitors, and market rules. For example, Danone's Danacol, a cholesterol-lowering yogurt, is a convergent offering, resulting from the comingling of the food and pharmaceutical industries. Danacol's competitor set includes players from both food and pharmaceutical industries. When a firm from one industry crosses over and enters a market in another industry, it creates a convergent offering. For example, Apple entered the cell phone handset industry through its iPhone, while remaining in the computing and music player industries. The convergent firm gains competitive advantage by being present in different industries with common technologies and sometimes, customers. Convergence can also be created when firms from different industries fulfill a customer need with different types of offerings. For example, Danone from the food industry offers a cholesterol-lowering yogurt, while Pfizer from the drug industry markets a cholesterol-reducing drug, Lipitor.

Competitive analysis will not be complete without a review of *coopetition* – the term used to describe the situation where competitors in one market cooperate in some other market(s) (Brandenburger and Nalebluff, 1996). Firms engaged in coopetition cooperate in those markets or situations where they believe that such cooperation will lead to favorable outcomes such as cost reduction, quality improvement, and profit increase. For example, Toyota and Peugeot, traditional competitors in many auto markets around the world, cooperated by sharing components costs for a new car they developed for the European market in 2005. A good competitive analysis could include assessment of coopetition opportunities and potential in different markets with different competitors. Such an analysis is typically done using game theoretic models.

Conclusion

Competitive analysis refers to the evaluation of strategies relating to a focal brand, product, SBU, or firm relative to current and potential competitors in its market or industry. Competitive analysis is undertaken before the formulation and development of positioning strategy, value propositions, and marketing mix strategies. It is also undertaken for a new brand or product, in particular, prior to its launch in a new market. Game theoretic models are useful tools

for competitive analysis. Competitive analysis in a new-product entry context involves a solid understanding of a framework that comprises the determinants of new-product entry strategy, its interrelationship with competitor response, and the drivers of competitor response. Competitive analysis can be enriched by including competitive benchmarking, convergence analysis, and coopetition analysis.

See also *competitor analysis; marketing strategy; marketing strategy models*

Bibliography

Bowman, D. and Gatignon, H. (1995) Determinants of competitor response time to a new product introduction. *Journal of Marketing Research*, **32** (1), 42–53.

Brandenburger, A. and Nalebluff, B. (1996) *Co-Opetition: A Revolution Mindset that Combines Competition and Cooperation: The Game Theory Strategy that's Changing the Game of Business*, Batam Doubleday Dell Publishing Group, Inc., New York.

Debruyne, M., Moenaert, R., Griffin, A. *et al.* (2002) The impact of new product launch strategies on competitive reaction in industrial markets. *The Journal of Product Innovation Management*, **19** (2), 159–170.

Gatignon, H., Anderson, E., and Helsen, K. (1989) Competitive reactions to market entries: explaining interfirm differences. *Journal of Marketing Research*, **26** (1), 44–55.

Gatignon, H., Robertson, T.S., and Fein, A.J. (1997) Incumbent defense strategies against new product entry. *International Journal of Research in Marketing*, **14** (2), 163–176.

Gruca, T., Kumar, K.R., and Sudarshan, D. (1992) An equilibrium analysis of defensive response to entry using a coupled response function model. *Marketing Science*, **11** (4), 348–358.

Hauser, J.R. and Shugan, S. (1983) Defensive marketing strategies. *Marketing Science*, **2**, 319–360.

Heil, O.P. and Walters, R.G. (1993) Explaining competitive reactions to new products: an empirical signaling study. *The Journal of Product Innovation Management*, **10** (1), 53–65.

Hultink, E.J. and Langerak, F. (2002) Launch decisions and competitive reactions: an exploratory market signaling study. *The Journal of Product Innovation Management*, **19** (3), 199–212.

Kalra, A., Rajiv, S., and Srinivasan, K. (1998) Response to competitive entry: a rationale for delayed defensive reaction. *Marketing Science*, **17** (4), 380–405.

Kerin, R., Varadarajan, R., and Peterson, R. (1992) First mover advantage: a synthesis and critique. *Journal of Marketing*, **56** (l), 33–52.

Kuester, S., Homburg, C., and Robertson, T.S. (1999) Retaliatory behavior to new product entry. *Journal of Marketing*, **63** (4), 90–106.

Kumar, K.R. and Sudarshan, D. (1988) Defensive marketing strategies: an equilibrium analysis based on decoupled response function models. *Management Science*, **34** (7), 805–815.

MacMillan, I., McCaffery, M.L., and Wijk, G.V. (1985) Competitors' responses to easily imitated new products – exploring commercial banking product introductions. *Strategic Management Journal*, **6** (1), 75–86.

Moorthy, K.S. (1985) Using game theory to model competition. *Journal of Marketing Research*, **22** (3), 262–282.

Porter, M.A. (1980) *Competitive Strategy*, Free Press, New York.

Porter, M.A. (1985) *Competitive Advantage*, Free Press, New York.

Robertson, T., Eliashberg, J., and Rymon, T. (1995) New product announcement signals and incumbent reactions. *Journal of Marketing*, **59** (3), 1–15.

Robinson, W. (1988) Marketing mix reactions to entry. *Marketing Science*, **7** (4), 368–385.

Shankar, V. (1997) Pioneers' marketing mix reactions to entry in different competitive games structures: theoretical analysis and empirical illustration. *Marketing Science*, **16** (4), 271–293.

Shankar, V. (1999) New product introduction and incumbent response strategies: their inter-relationship and the role of multimarket contact. *Journal of Marketing Research*, **36** (3), 327–344.

Shankar, V. (2006) Proactive and reactive product line strategies: asymmetries between market leaders and followers. *Management Science*, **52** (2), 276–292.

Shankar, V., Ancarani, F., and Costabile, M. (2010) Strategically Leveraging Convergence. Working Paper, Texas A&M University, College Station.

Shankar, V., Carpenter, G.S., and Krishnamurthi, L. (1998) Late mover advantage: how innovative late entrants outsell pioneers. *Journal of Marketing Research*, **35** (1), 54–70.

Shankar, V., Carpenter, G.S., and Krishnamurthi, L. (1999) The advantages of entering in the growth stage of the product life cycle: an empirical analysis. *Journal of Marketing Research*, **36** (2), 269–276.

Varadarajan, R.P., Yadav, M., and Shankar, V. (2008) First-mover advantage in the internet-enabled environment: a conceptual framework and propositions. *Journal of Academy of Marketing Science*, **36** (3), 293–308.

competitor analysis

Venkatesh Shankar

Competitor analysis refers to the strategic evaluation of the strengths and weaknesses of current and potential competitors for a focal brand in its market(s) or for a focal strategic business unit (SBU) or firm in their industry(ies). Competitor analysis can be done at the product-market level, the SBU level, or the firm level. The main purpose of competitor analysis is to develop business and marketing strategies that are based on sustainable competitive advantage. The dimensions of strengths and weaknesses include financial resources, human resources, core capabilities and competencies, order of entry advantage or disadvantage, manufacturing expertise, research and development (R&D) expertise, brand equity, distribution clout, pricing power, promotional capability, and execution ability. Each current and potential competitor is assessed on these dimensions and a profile of each competitor is developed. An overall assessment is made on who the key competitors are and what critical strengths and weaknesses they bring to bear on the market. Such an assessment helps brands and firms anticipate future competitor actions and reactions and exploit any weaknesses or gaps in the market or the industry. Thus, competitor analysis provides a critical input to the development of a firm's competitive marketing strategy.

At the firm and SBU levels, competitor analysis involves the assessment of all relevant firms in the relevant industries. When applied at the brand level, competitor analysis is useful in formulating brand-positioning strategy. Customer perceptions of each brand's strengths and weaknesses in a given product market provide the inputs for deriving a perceptual map for that market. Using this information and data on the location and size of customer preferences, a brand can develop its positioning strategy. Competitor analysis helps brands differentiate themselves from other brands.

A good competitor analysis should include assessments of both current and potential future competitors. Future competitors include marketers of substitute products and those with alternative technologies capable of satisfying customer needs in the relevant markets or industries. Competitors are typically identified first in the core product market and then in the adjacent or peripheral product markets. The most direct or relevant competitors can be identified on the basis of an analysis of the market structure, which includes such considerations as similarity of served needs, substitutability, cross-price elasticity, switching costs, and perceptual positions. With the rapid global growth in the penetration of the Internet, the competitor sets for most brands and firms have broadened to include players from different product markets and countries. Furthermore, a firm's competitor in one market could be its collaborator in another market, in particular, in the hi-tech industry. A sound competitor analysis should incorporate the implications of such possibilities. The competitor set for any brand or business unit or firm is dynamic as new players enter its markets and existing players leave those markets. Therefore, competitor analysis is performed at intervals of a year or less.

Competitor analysis is a useful tool for deciding whether to launch a new product. By combining competitor analysis with an analysis of own strengths and weaknesses and of external opportunities and threats, firms can better decide whether to introduce a new product. Firms typically decide to launch products when their strengths relative to their competitors outweigh their relative weaknesses in markets where opportunities surmount threats.

The term, *competitor analysis*, is used interchangeably with another term, *competitive analysis*, although competitive analysis typically refers to the assessment of a business situation with a view to develop a sound business strategy or marketing strategy relative to competitors. Competitor analysis provides an important basis for the development of such a strategy. Competitor analysis is an essential element of any marketing plan. In a typical marketing plan, it is a part of situational analysis, which is followed by marketing objectives and marketing strategy.

See also *competitive analysis; marketing strategy; marketing strategy models*

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Bibliography

Brandenburger, A. and Nalebluff, B. (1996) *Co-Opetition: A Revolution Mindset that Combines Competition and Cooperation: The Game Theory Strategy that's Changing the Game of Business*, Batam Doubleday Dell Publishing Group, Inc., New York.

Gatignon, H., Anderson, E., and Helsen, K. (1989) Competitive reactions to market entries: explaining interfirm differences. *Journal of Marketing Research*, **26** (1), 44–55.

Hauser, J.R. and Shugan, S. (1983) Defensive marketing strategies. *Marketing Science*, **2**, 319–360.

Heil, O.P. and Walters, R.G. (1993) Explaining competitive reactions to new products: an empirical signaling study. *The Journal of Product Innovation Management*, **10** (1), 53–65.

Kerin, R., Varadarajan, R., and Peterson, R. (1992) First mover advantage: a synthesis and critique. *Journal of Marketing*, **56** (l), 33–52.

MacMillan, I., McCaffery, M.L., and Wijk, G.V. (1985) Competitors' responses to easily imitated new products – exploring commercial banking product introductions. *Strategic Management Journal*, **6** (1), 75–86.

Moorthy, K.S. (1985) Using game theory to model competition. *Journal of Marketing Research*, **22** (3), 262–282.

Porter, M.A. (1980) *Competitive Strategy*, Free Press, New York.

Porter, M.A. (1985) *Competitive Advantage*, Free Press, New York.

Shankar, V. (1997) Pioneers' marketing mix reactions to entry in different competitive games structures: theoretical analysis and empirical illustration. *Marketing Science*, **16** (4), 271–293.

Shankar, V. (1999) New product introduction and incumbent response strategies: their inter-relationship and the role of multimarket contact. *Journal of Marketing Research*, **36** (3), 327–344.

Shankar, V. (2006) Proactive and reactive product line strategies: asymmetries between market leaders and followers. *Management Science*, **52** (2), 276–292.

Shankar, V., Ancarani, F., and Costabile, M. (2010) Strategically Leveraging Convergence. Working Paper, Texas A&M University, College Station.

Shankar, V., Carpenter, G.S., and Krishnamurthi, L. (1998) Late mover advantage: how innovative late entrants outsell pioneers. *Journal of Marketing Research*, **35** (1), 54–70.

Shankar, V., Carpenter, G.S., and Krishnamurthi, L. (1999) The advantages of entering in the growth stage of the product life cycle: an empirical analysis. *Journal of Marketing Research*, **36** (2), 269–276.

Varadarajan, R.P., Yadav, M., and Shankar, V. (2008) First-mover advantage in the internet-enabled environment: a conceptual framework and propositions. *Journal of Academy of Marketing Science*, **36** (3), 293–308.

customer analysis

Barry J. Babin and Eric G. Harris

Few would argue with the notion that customers hold a key to success for businesses in competitive industries. Customers are the subset of consumers who choose to interact meaningfully with a particular business entity. Businesses, therefore, are challenged to convert consumers into customers and once they have done so, to retain them as customers. Value is the key tool that businesses use to accomplish this feat. Thus, businesses need intelligence about consumers if they are to make reasoned strategic and tactical decisions; these decisions are all directed at value creation. Customer analysis, then, is the process of generating intelligence from information collected about consumers in the form of data. Eventually, managers make this intelligence actionable through informed business strategy and tactics.

Theory plays a large role in customer analysis. In particular, analysts need to know what needs to be studied and how to study it. The "what" needs to be studied question is addressed through the accumulated body of theory dealing with consumer behavior. This theory in itself is largely derived from other disciplines including psychology, social psychology, anthropology, and economics. The "how to" study question is addressed through the accumulated body of knowledge dealing with consumer research. In a very real way then, consumer theory is firmly rooted in the philosophy of science and the method of learning about consumers is structured around the scientific method. Not only does consumer research reveal ways to gather information about and from consumers but it also is increasingly focused on ways of analyzing this data. This analysis brings the data to life by converting raw information into intelligence.

The purpose of this article is to provide a basic overview of customer analysis. We offer a perspective to customer analysis based on two frameworks, one, a framework for understanding consumer value, the other, a framework for generating consumer intelligence. The article illustrates how research and analysis vary on the basis of the particular aspect of consumer behavior being studied. No encyclopedia of marketing would be complete without insight into studying consumers and this article is intended to fill this role.

Customer Value

Any attempt to analyze consumer behavior must begin with recognition of the importance of value. Value is the sine qua non of consumer behavior. In fact, according to Peter Drucker, the marketing concept should be focused on the creation of value for consumers (Uslay, Morgan, and Sheth, 2008). Furthermore, the American Marketing Association acknowledges the key role of value in its very definition of the field of marketing:

> "Marketing is the activity, set of institutions, and processes for creating, communicating, delivering, and exchanging offerings that have value for customers, clients, partners, and society at large." (AMA, 2007, http://www.marketpower.com/About AMA/Pages/Definitionof Marketing.aspx).

While there have been a number of conceptualizations offered in the literature regarding the value construct, most can be summed up in one simple definition; *value* represents a personal assessment of the net worth obtained from an activity (Babin and Harris, 2009). As such, value amounts to the gratification that one receives from consumption or the total net benefit of service received. Importantly, both actors in an exchange give and receive value. Marketers receive value when they obtain resources, most often financial resources, in exchange for benefits derived from using or enjoying a product offering. Consumers receive value after a purchase has taken place through the process of consumption. For consumers, then, a simplified way of viewing value is through a function that highlights what consumers give up and what they get from an exchange. This follows a traditional view of value as "what I get versus what I give" (Zeithaml, 1988). This function is illustrated in Figure 1.

Consumers typically get benefits such as quality, convenience, and emotions from their exchanges with marketers. Likewise, they give up resources such as time, effort, money, and energy. Although this is a rather simple way of

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

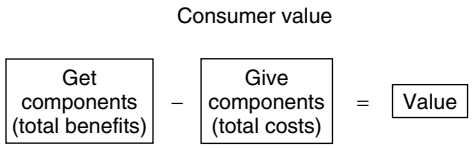


Figure 1 Consumer value.

illustrating the value concept, it does illustrate the important "give and get" components.

Dimensionality of Value

An important issue with regard to consumer value is the dimensionality of the construct. Although the dimensionality of value has received scant research attention, one useful typology suggests that two main components of value help capture the construct *utilitarian value* and *hedonic value* (Babin, Darden, and Griffin, 1994). A consumer realizes the utilitarian value when he/she purchases a product that, when consumed, helps solve specific problems. For example, when a consumer buys something as mundane as a stapler, there is a specific problem that is to be solved. Attaching papers together creates greater organization and facilitates the user's work. In this way, utilitarian value is thought of as a "means to an end" (Babin, Darden, and Griffin, 1994). In other words, by buying a stapler the consumer is able to solve a problem that is encountered frequently. The stapler, in and of itself, provides a solution to a common, tangible problem, though owning it provides no value. That is, until it is used, it is relatively useless.

Hedonic value is focused on the immediate gratification that results from some consumption activity (Babin and Harris, 2009). This value is associated with the act of consumption and the emotions that come along with it. Little hedonic value is obtained from the act of stapling pieces of paper together, however, a great deal of hedonic value is found in a visit to an amusement park, a sporting event, an art museum, or a fine dining establishment. Consuming fine champagne is a multisensory experience that epitomizes hedonic value, whereas utilitarian value is influenced by evaluation of consumption components such as convenience and efficiency. Hedonic value is driven by the emotions and feelings that emerge during the act of consumption (Babin and James, 2010). One seldom goes to a theater production to be educated. Generally, the theater production is consumed primarily because the time invested as an audience member is gratifying immediately.

Consumer Value Framework: The "What" of Consumer Analysis

As a general statement, then, customer analysis must begin with an appreciation of the importance of value, its dimensionality, its role in consumer behavior, as well as the myriad of variables that influence the construct. Some theoretical orientation is necessary to narrow down the range of potential phenomena that might be useful in explanations. One constructive framework for understanding the many variables that can influence the value derived from consumption is the "consumer value framework (CVF)" (Babin and Harris, 2009). This framework is illustrated in Figure 2. The framework addresses the "what" question of consumer analysis by providing a snapshot of important variables to consider in any consumer research effort.

Figure 2 indicates that numerous variables work to influence consumption-related behaviors, and ultimately, the value that consumers receive from products, services, and experiences. Internal influencers are derived directly from the field of psychology. Here, we find topics such as learning, perception, memory processes, consumer information processing, categorization, attitude, and personality. All of these concepts are crucial for studying consumer behavior. External influences, long recognized in the field of social psychology, include topics such as acculturation/enculturation, culture, reference groups, social class, and family influence. Situational influencers are also important to consider here. These include topics such as atmospherics, time pressure, and conditions.

The relationship between the marketer and the consumer also plays an important part in all of consumer behavior and analysis. Firms commonly focus on customer relationship management (CRM) techniques to ensure that long-lasting relationships exist between company and customer. Important topics here include customer satisfaction/dissatisfaction,

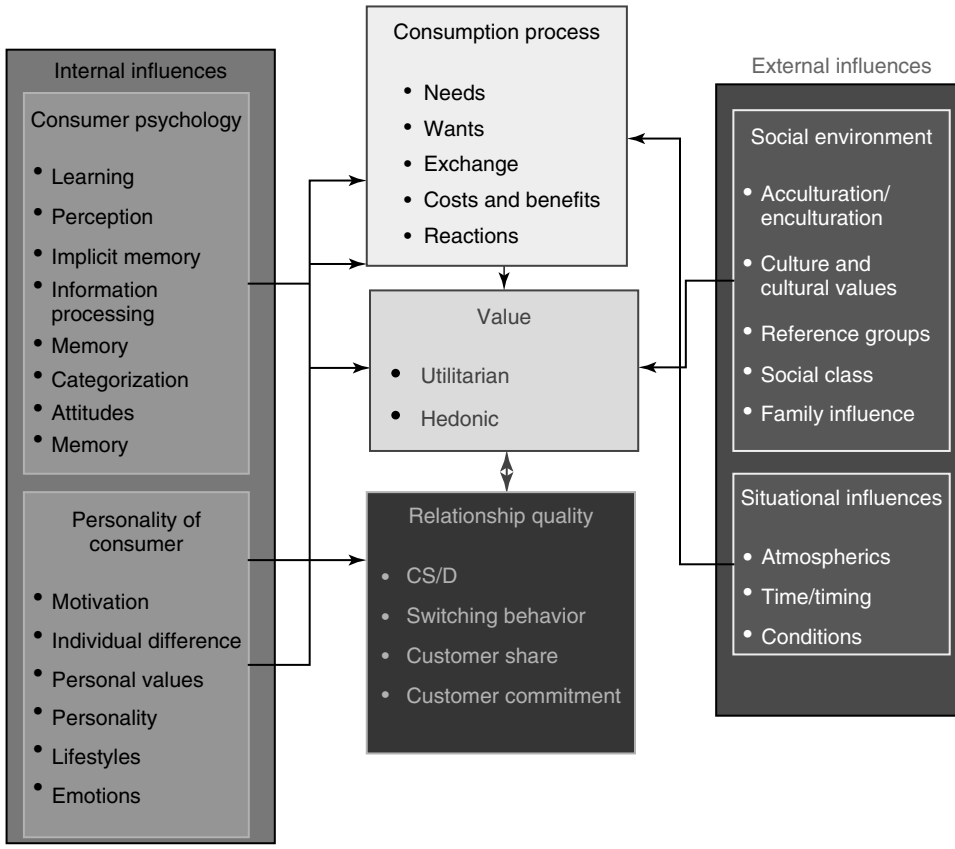


Figure 2 Consumer value framework (see Babin and Harris, 2009).

consumer switching behavior, customer "share," and customer commitment.

It is important to emphasize that much research attention has been given to the customer satisfaction/dissatisfaction concept at the expense of customer value. As Babin and James 2010 discuss, however, a great deal is to be gained by focusing on customer value. In fact, many highly successful marketers (e.g., Walmart) have shown that delivering customer value is a major part of a successful marketing strategy. Of critical importance here is the realization that customer value drives customer loyalty in a way in which customer satisfaction does not.

Customer Analysis in Practice

While the CVF provides an integrated view of the "what" to study in consumer analysis, the question of "how" to perform consumer analysis is answered by following scientific research principles applied specifically to customer analysis (see Figure 3). Although the research process is not the primary focus of the current work, we can say that research follows a scientific process aimed at uncovering truths – in this case, truths about consumers (Zikmund and Babin, 2010). Customer analysis focuses on uncovering the myriad of factors that influence customer value. This process can occur in both applied and basic research settings, that is, customer analysis can be applied to very specific situations and/or problems (as when a company analyzes its customers in order to increase sales), or it can be applied by academicians who attempt to gain better understanding of consumer psychology and how it impacts buying decisions, and ultimately, value.

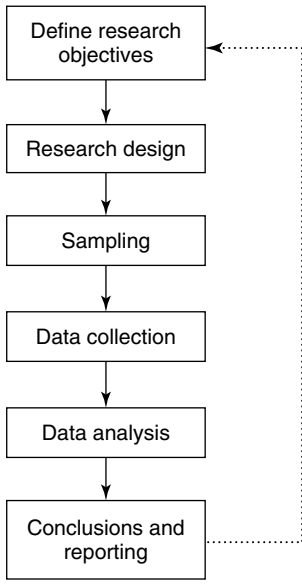


Figure 3 A framework for research aimed at customer analysis.

Customer analysis, therefore, involves research. As a result, it follows basic research principles and the basic steps of the scientific method: (i) defining the analysis objectives; (ii) planning an analysis research design; (iii) planning a sample; (iv) collecting data; (v) analyzing the data; and (vi) formulating conclusions. Although it is quite common that most marketing research efforts focus on relatively large samples, customer analysis is often applied to individual consumers. Of course, this is dependent upon the purpose of the research.

The three basic categories of consumer research designs are *exploratory, descriptive, and causal research*. *Exploratory research* is performed when the researcher knows very little about the particular subject matter or situation. The goal here is to simply gain an understanding of the consumer behavior phenomena with the expectation that more conclusive research is likely to follow. Popular techniques for exploratory research studies include pilot studies, case studies, and experience surveys. Interpretive research is often used as an exploratory research tool. An interpretive approach can involve ethnography or phenomenology. In either case, the researcher takes on a more involved role in observing and/or interacting with research respondents. The data can take the form of interview text, a story completion, a collage, or other observational field notes. An interpretive researcher then derives meaning from this data and the meaning can be useful in establishing a potential explanation. *Descriptive research* is undertaken when the researcher has a basic idea of the subject matter and simply wants to describe the characteristics of people, situations, organizations, or environments. Surveys, secondary studies, and observation techniques are popular tools for descriptive research. Common statistical tools include cross-classification, independent sample t-tests, multiple regression, and factor analysis. *Causal research* is performed when the research has the purpose of establishing "cause and effect" relationships between variables. Experimental designs are popular tools for causal research. Experimental research often involves statistical analysis using some form of an analysis of variance (ANOVA) model. Experimental designs are particularly common in basic consumer research such as that published in the *Journal of Consumer Research*.

Customer analysis approaches. Customer analysis can be based on results from either *qualitative* or *quantitative* research. Qualitative approaches are often interpretive and focus heavily on researcher interpretation with little or no regard for numerical measurement of consumer phenomena, whereas quantitative research techniques rely heavily on numerical measurement of research observations. Qualitative techniques are often used with relatively small sample sizes. These approaches include techniques such as *phenomenology* (the interpretation of the lived experience), *ethnography* (a technique that requires the researcher to be immerged in the research context itself), and *focus groups* (unstructured interviews of a small group of consumers). Such techniques are said to be *researcher dependent* because they rely heavily on the interpretations of the observer. These techniques have gained in popularity among consumer researchers and are utilized regularly in customer analysis. In contrast, quantitative approaches involve techniques like

survey designs, experimentation, and statistical analysis. Quantitative approaches often use relatively large sample sizes. Regardless of the research design that is utilized, researchers must work to ensure that the results obtained are both valid and reliable. There are several techniques available for assessing the validity and reliability of customer analysis studies, though these techniques are beyond the scope of the current work. Qualitative approaches such as these are very commonly employed to develop explanations of the way customers might behave.

Data in customer analysis. There are many ways in which data can be collected for customer analysis. Survey methods generally require that respondents complete mail, e-mail, or door-to-door surveys. The survey procedure can sometimes require the use of a trained interviewer or research assistant who helps record consumer responses. Or, in some circumstances, respondents complete the surveys themselves. Data can also be collected by trained observers who simply observe consumer behavior. Yet another way of collecting data is by using some type of mechanical equipment such as video recorders or sophisticated psychological equipment such as *psychogalvanometers* or *voice-pitch analyzers*. We emphasize once again, however, that the focus of customer analysis is often on individual customer responses. As such, qualitative techniques such as depth interviews are extremely applicable. Customer data gathered using a qualitative approach lends itself more directly to interpretive scrutiny than to statistical analysis.

Data analysis. Data analysis is a critical phase of customer analysis. In this phase, the raw data that are collected in the study are transformed via some type of analytical technique into meaningful information. Some refer to this step as *data transformation*. The selection of analytic technique depends on the specific research design.

Qualitative or interpretative designs often require some type of categorization or theme generation. In this step, the researcher attempts to categorize data in either existing or "emergent" categories. In doing so, the researcher must consider the context from which responses are taken. For example, a researcher may focus on the meanings that consumers attach to possessions that are handed down from generation to generation. To do so, they would rely heavily on interviews and life stories of family members in order to gain an understanding of the meaning of such possessions. This would take place in the absence of consumer surveys or other quantitative approaches. A great deal of training is necessary for a researcher to effectively interpret qualitative data.

There are numerous quantitative analysis techniques that are available to the consumer researcher. As mentioned earlier, survey and experimental designs lend themselves well to data analysis tools such as *regression analysis, ANOVA, correlation analysis, and t-tests*. However, researchers must often utilize more sophisticated multivariate data analysis techniques such as *cluster analysis, discriminant analysis, canonical correlation, conjoint analysis, and structural equations model* (*SEM*). Although the details of these techniques are beyond the scope of the current article, it is worth noting that these techniques are used commonly in consumer analysis. Several software packages are available, including *SPSS*, *SAS*, *LISREL*, *Mplus*, and *AMOS*. Quantitative techniques generally require a great deal of training in specific statistical theory. Regardless of the complexity of the technique, however, the goal of most quantitative research remains the same: to obtain information from the sample data that can be inferred to the population of interest.

Drawing conclusions. Transforming data into usable information is an important part of the customer analysis process. Regardless of the complexity (or simplicity) of the study, useful information is critical. In fact, communicating research results is one of the most important phases of the entire research process (Zikmund and Babin, 2010). Of course, the audience of the research study must be considered. For applied research studies, the audience generally consists of managers who are seeking to make business decisions that are based on the information

Table 1 Examples of research aimed at consumer analysis.

Area of CB Involved	*Research Question*	*Research Design*	*Research Description*	*Results*	*Reference Citation*
Internal influences – motivation value	What is the nature of consumer motivation in pet consumption?	Exploratory	Interpretive research involving ethnography of show dog owners.	Specific insights into motivations of dog-related marketing.	Bettany and Daly (2008)
Internal influences – personality of consumer Consumption process	What are the lifestyle characteristics distinguishing heavy online shoppers versus online shopping avoiders?	Descriptive	A survey using a probability sample representing US consumers.	Online shoppers are relatively young, wealthy, and live a lifestyle highly oriented around the Internet.	Swinyard and Smith (2003)
External influences – family influence	How does the degree of family communication influence the transmission of brand attitudes from parents to children?	Descriptive	A survey of elementary school children in Taiwan.	Brand attitudes of Taiwanese children are influenced by parents and these effects are influenced by social and concept orientations.	Hsieh, Chiu, and Lin (2006)
External influences – culture and situation value	How do Asian consumers from different countries react to different levels of service performance?	Causal	A lab experiment manipulating characteristics of the service environment (atmosphere).	Chinese subjects display greater variance in satisfaction than do Korean or Japanese subjects.	Ueltschy *et al.* (2009)
External influences – social class	What is the role of disempowerment as a central phenomenon of social class?	Exploratory	A series of in-depth interviews of workers in Sydney, Australia.	Relative empowerment contributes to distinctive financial approaches to decision making	Henry (2005)
Situational influences on decision making	How do repeated incident exposures to features of the everyday environment influence product evaluation and choice?	Causal	A combination of quasi and lab experiments utilizing samples of US consumers.	The prevalence of perceptually and conceptually related stimuli can shape product judgment and decision making	Berger and Fitzsimons (2008)

Table 1 (*continued*).

Area of CB Involved	*Research Question*	*Research Design*	*Research Description*	*Results*	*Reference Citation*
Influence of atmospherics on decision making	When and how does the meaning of music influence perceptions of an advertised product?	Causal	Lab experiments manipulating need for cognition and processing intensity, message format, and type of music.	Nonintensive processors are insensitive to the meaning of music but intensive processors base perceptions on meaning.	Zhu and Meyers-Levy (2005)
Postconsumption processes	What is the relationship between hedonic and utilitarian benefits and postconsumption feelings of delight and satisfaction?	Causal	A lab experiment manipulating product design benefits and consumption experience.	A product that meets or exceeds utilitarian expectations evokes satisfaction, whereas a product that meets or exceeds hedonic expectations evokes delight.	Chitturi, Raghunathan, and Mahajan (2008)
Decision making and value	What are elements that translate into value from the purchase of high-tech products by business customers?	Exploratory	Interpretive approach based on data from in-depth interviews of business customers.	Business perspective on what are important elements varies on the basis of the role one plays in the decision process.	Lindgreen *et al.* (2009)
Consumer cognition and decision making	How does the inclusion of functional health claims on food products alter perceptions of value (i.e., choice)?	Causal	A lab experiment manipulating product labeling – specifically functional health claims.	Health conscious consumers are most likely to choose something with a functional claim – even when presented with conflicting information.	Naylor, Droms, and Haws (2009)

provided. For basic research studies, the audience largely consists of consumer researchers who are seeking to learn more about consumer behavior so that they can advance consumer research theory.

Example Applications for Customer Analysis

Table 1 provides a brief overview of published examples of the types of analytical approaches

that have been described . Each study involves drawing generalizations about consumers with varying degrees of certainty. Some areas of consumer behavior lend themselves better to certain types of research designs. Studies of consumer motivation sometimes prove difficult for survey research or even causal research. Manipulating motivation can be troublesome. Thus, interpretive research is often employed. These techniques are often aimed at relatively small samples of consumers. Bettany and Daly (2008) employ this approach in studying consumer motivations for pet-related consumption activities. In contrast, studies involving individual difference characteristics like demographics, personality, or lifestyle, lend themselves very well to descriptive research employing consumer surveys, often with large sample sizes. The Swinyard and Smith (2003) study analyzes characteristics of online shoppers and typifies this approach. Finally, studies involving consumer perception, atmospherics, and decision processes lend themselves well to experimental designs. This is because the characteristics of different marketing stimuli such as advertising copy or even physical elements like odor can be manipulated easily in laboratory environments (see Ueltschy *et al.*, 2009). Basic (or academic) research analyzing results from studies effectively matching the analytical approach to the research question about consumers has contributed and will continue to contribute to the development of consumer behavior theory and to effective analysis of customers by practitioners.

Consumer analysts also recognize that the term customer does not always involve an individual family member or even a household. Businesses are customers too and the contribution of business-to-business spending to the GDP of most developed nations is approximately equal to that of business-to-consumer spending. Thus, a significant amount of customer analysis addresses business consumers. Lindgreen *et al.* (2009), for example, provide a typical business customer analysis, in this case, employing a qualitative approach, in which they explore the key elements that make a high-tech product more valuable in the eyes of the business customer. Likewise, customer research is not always motivated by a specific marketing problem for a business. The government is also interested in customer analysis. Each public policy decision that involves a restriction in the marketplace should be informed by research. Food product labeling is regulated in every developed country. Thus, a significant amount of consumer research addresses the ways consumers react to different types of product claims and different ways to present health-related claims. Researchers, for example, analyze the way customers react to claims that a product provides some health function such as aiding digestion (Naylor, Droms, and Haws, 2009). In this case, a causal design allows analysis aimed at more informed government decisions on labeling.

No matter how the research is generated, customer analysis provides critical information for theory development, businesses, and governments. Without valid analysis, decision makers are left to make decisions based solely on intuition, politics, or pure guess work. For instance, decisions about what consumer segments to appeal to involve analysis matching the personality of a consumer to the personality of a brand; decisions about how to communicate nutrition information appropriately involve analysis of psychological processes; and decisions about environmental design involve analysis of key situational and cultural variables. Intelligence is the key to reducing the risk of decision making. In the end, valuable outcomes for consumers can best be delivered when decision makers have some idea of the processes by which value can be created.

See also *brand value; competitor analysis; customer lifetime value (CLV); customer solutions; hedonism: gratification is value; marketing research process; structural equation modeling; survey research; thinking deeper about customer experience; value co-creation*

Bibliography

AMA (2007) http://www.marketpower.com/AboutAMA/Pages/Definitionof Marketing.aspx (accessed July 20).

Babin, B.J., Darden, W.R., and Griffin, M. (1994) Work and/or fun: measuring hedonic and utilitarian shopping value. *Journal of Consumer Reasearch*, **20**, 644–656.

Babin, B.J. and Harris, E.G. (2009) *CB. A Value-Based Approach*, South-Western Cengage Learning, Mason.

Babin, B.J. and James, K.W. (2010) A brief retrospective and introspective on value. *European Business Review*, forthcoming.

Berger, J. and Fitzsimons, G. (2008) Dogs on the street, pumas on your feet: how cues in the environment influence product evaluation and choice. *Journal of Marketing Research*, **45** (1), 1–14.

Bettany, S. and Daly, R. (2008) Figuring companion-species consumption: a multi-site ethnography of the post-canine Afghan Hound. *Journal of Business Research*, **61**, 404–418.

Chitturi, R., Raghunathan, R., and Mahajan, V. (2008) Delight by design: the role of hedonic versus utilitarian benefits. *Journal of Marketing*, **72**, 48–63.

Henry, P. (2005) Social class, market situation, and consumer metaphors of (Dis) empowerment. *Journal of Consumer Research*, **31**, 766–7780.

Hsieh, Y.-C., Chiu, H.-C., and Lin, C.-C. (2006) Family communication and parental influence on children's brand attitudes. *Journal of Business Research*, **59**, 1079–1086.

Lindgreen, A., Antioco, M., Palmer, R., and Heesch, T. (2009) High-tech, innovative products: identifying and meeting business customers' value needs. *Journal of Business and Industrial Marketing*, **24** (3), 182–197.

Naylor, R.W., Droms, C.M., and Haws, K.L. (2009) Eating with a purpose: consumer response to functional food health claims in conflicting versus complementary information environments. *Journal of Public Policy and Marketing*, **28**, 221–233.

Swinyard, W.R. and Smith, S.M. (2003) Why people (Don't) shop online: a lifestyle study of the internet consumer. *Psychology and Marketing*, **20** (7), 567–597.

Ueltschy, L.C., Laroche, M., Zhang, M. *et al*. (2009) Is there really an Asian connection? Professional service quality perceptions and customer satisfaction. *Journal of Business Research*, **62**, 972–979.

Uslay, C., Morgan, R.E., and Sheth, J.N. (2008) Peter drucker on marketing: an exploration of five tenets. *Journal of the Academy of Marketing Science*, **37**, 47–60.

Zeithaml, V.A. (1988) Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence. *Journal of Marketing*, **52**, 2–22.

Zhu, R. and Meyers-Levy, J. (2005) Distinguishing between the meanings of music: when background music affects product perceptions. *Journal of Marketing Research*, **42** (3), 333–345.

Zikmund, W. and Babin, B.J. (2010) *Exploring Marketing Research*, 10th edn, South-Western Cengage Learning, Mason.

thinking deeper about customer experience

Jerry Olson and Gerald Zaltman

In recent years, marketers have become aware of the importance of customer experience (Carbone, 2004). Service companies such as theme parks, restaurants, and airlines quickly recognized that an emotional experience is the primary benefit they provide. Somewhat belatedly, product-oriented companies realized that tangible products also provide customers with important experiences.

We propose that a person's thoughts and feelings about an event are central to customer experience. Table 1 shows that marketers have not always treated customer experience in that way.

In more than a few cases, observations of "what happened" have been treated as experience. Marketers can learn "what happened" through secondary sources (How many people repurchased brand A within six weeks?) or ethnographic observation (How many brands do consumers examine in the cereal aisle?) (*see* DESCRIPTIVE RESEARCH; PERSONAL OBSERVATION). Although such descriptions can be useful, they are not true experience. Rather, experience resides in a customer's interpretation of "what happened." When marketers do not know how customers understand "what happened," they must imagine the meanings and feelings that constitute experience. That intuition is often wrong and can lead to costly mistakes.

Identifying customers' feelings of "what happened" as an event occurs does reflect true experience, but measuring such phenomenological reactions is difficult. Consider how we could measure customers' real-time thoughts and feelings during the course of a grocery shopping trip. We are able to capture physiological reactions (e.g., galvanic skin response, heart rate, or perhaps eye fixations) for relatively simple behaviors such as reactions to an ad. However, physiological measures do not reveal customers' cognitive and emotional interpretations. Moreover, we know that quite different feelings can produce similar physiological signatures. For more complex experiences such as a day at a theme park, it is easier to obtain customers' self-reported thoughts and feelings immediately following the behaviors of interest. However, these "direct" measures can access only phenomenological reactions of which the customer is consciously aware, and we know that significant aspects of customers' experience will be unconscious.

Perhaps the best indicators of customers' experience access memories of "what happened," since memories reflect prior interpretations by the customer (*see* CONSUMER MEMORY PROCESSES). Memory-based measures are influenced by many factors, including the amount of time since "what happened," the number of prior recalls of "what happened," and the type of questioning. Direct requests to remember "what happened" are likely to yield relatively shallow accounts of experience. To identify the deeper aspects of experience, including customers' mostly unconscious interpretations of "what happened," we need more sophisticated measures.

The Zaltman Metaphor Elicitation Technique (ZMET) uses metaphor to help customers access and describe the deeper, often unconscious meanings associated with an experience (Zaltman, 2003). In a ZMET study, customers are asked to find pictures (sometimes, objects) that represent their thoughts and feelings about an experience of interest. Each picture is a metaphor for some relevant aspect of the customer's experience; often an unconscious meaning, emotion, or orientation that cannot be accessed by direct questioning. In projecting their thoughts and feelings onto the pictures, customers tell stories that reveal the interpreted meanings of their experience, often in rich detail (*see* PROJECTIVE TECHNIQUES).

For example, consider a ZMET study in which we asked theatergoers to find images that express their Broadway theater experience. The stories these customers told revealed the conscious and unconscious feelings and meanings of their remembered Broadway experiences. A key unconscious metaphor was Container, indicated by customers' stories about the various "places" where they experienced Broadway theater. As customers moved in and out of the various "Containers of Broadway meaning," they experienced different thoughts and emotions. For instance, theatergoers described

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Three types or levels of experience.

Type of Experience	*Interpretation*	*Possible Measures*	*Example*
Observed behavior	What happened	An external observer's account of customers' behaviors in a situation as captured by ethnography or video recording.	Susan waited in line at Starbucks for 3.45 minutes.
Phenomenological experience	The feeling of "what happened," as it occurred	Physiological measures taken during the situation; or immediate self-reported feelings.	Susan felt only slightly frustrated by her wait as she was distracted by the product displays around her.
Remembered experience	The meaning of what happened, as stored in memory	Retrospective interpretation of one's behavior in a situation. This process blends three things: the memory of what happened, the memory of the feelings that were experienced, and the currently activated mental frame.	Over lunch, Susan told a friend about her recent visit to Starbucks. Her feelings about waiting in line to be served were not a relevant aspect of her experience; rather she described two products that caught her interest during that time.

feelings of excitement and anticipation as they entered the hustle and bustle of the Broadway area in New York. When customers entered the "different world" of the theater building, they imagined theater customers of bygone days and felt a connection with them. The play also was framed as a "container" into which theatergoers became immersed. Often, they became "lost in" the story line as they "explored" how the play is relevant to issues in their own lives. After leaving the theater, some customers entered into extensive discussions with friends about their feelings and interpretations, the final "container" of their Broadway theater experience.

We propose that marketers develop and apply measurement tools like ZMET that can uncover the deep and rich aspects of customer experience. Gaining a deeper understanding of how customers think and feel about "what happened" will help marketers develop more effective ways to improve the experiences of their customers (*see* CUSTOMER SATISFACTION/DISSATISFACTION; CUSTOMER-SATISFACTION RESEARCH).

Bibliography

Carbone, L.P. (2004) *Clued, in How to Keep Customers Coming Back Again and Again*, Upper Saddle River, Financial Times Prentice Hall.

Zaltman, G. (2003) *How Customers Think: Essential Insights into the Mind of the Market*, Harvard Business School Press, Boston.

customer lifetime value (CLV)

Katherine N. Lemon and Loren J. Lemon

Customer lifetime value (CLV) is defined as the net profit from a customer (customer margin less the costs of acquiring and serving the customer) over the lifetime of the customer's relationship with the firm, discounted to the present. One of the key objectives of CLV is to assist in the estimation of how much a customer/customer segment (new or existing) is worth to a firm over the lifetime of the customer's relationship with the firm in current terms (e.g., today's dollars). While the terminology may vary from model to model, the essential elements of CLV estimation consist of the following:

- Acquisition cost
- Time interval
- Net profit
- Retention rate
- Discount rate

To understand CLV, it is important to understand the definitions of each element of CLV.

First, acquisition cost is equal to the costs the firm incurs in first securing the customer less any revenues derived by the firm from the customer at the time of acquisition. Second, time interval is the time periods (usually years) and overall time horizon selected for a given CLV estimation. While this may be infinite, usually a specified interval is selected based on the firm's general "life expectancy" of its relationships with the customer(s) being analyzed. An appropriate time interval for most CLV estimates is three to five years. Third, net profit is the expected future profits to be derived from the customer(s) being analyzed for the selected time interval. This is often considered to be the "net margin" from such transactions or the recurring revenue derived from the customer less any recurring costs in servicing the customer. These costs should ideally be directly associated with the customer. Fourth, retention rate is the expected probability that the firm will retain the customer as a customer during the selected time interval. Finally, discount rate is the rate used in calculating the "net present value" of the future profits derived from the customer. The usual reference point in selecting the discount rate is the firm's weighted average cost of capital.

The basic model of CLV is presented below. CLV is typically expressed as the following mathematical formula in Equation 1:

$$\sum_{t=0}^{T} m \frac{r^t}{(1+d)^t} - AC \qquad (1)$$

where

AC = Acquisition costs,
t = Time interval,
d = Discount rate,
m = Net profit,
r = Retention rate.

When the following specific assumptions are satisfied, the equation simplifies to Equation 2 below:

- recurring net profit (m) is constant over the life of the customer,
- retention rate (r) is constant over the life of the customer,
- discount rate (d) is constant over the life of the customer.

$$CLV = \frac{m}{1-r+d} - AC \qquad (2)$$

It is important to note that CLV estimations can be made for a specific customer, specific customer segment, or a generalized calculation based on the firm's generalized experience for its customers. For decisions based upon CLV to be actionable, the customer segment estimated should be as homogeneous as possible. CLV is useful in assisting a firm in making a determination of whether any new initiatives by the firm are worthwhile (Berger and Nasr, 1998). For example, initiatives focusing on acquiring new customers can be evaluated to determine whether the costs of such new initiates may be worth it given the potential expected future profits from such customers

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

(*see* CUSTOMER RELATIONSHIP MANAGEMENT). Initiatives focusing on current customer (either to increase the revenues from existing customer or improving the retention rate for such customers) can be evaluated by first estimating the CLV for such customers prior to the initiative and then reestimating the expected CLV with the new initiatives.

CLV can also be used to determine which customer segments to focus on, which customers may not be worth retaining, and where sales and marketing efforts and other firm resources may best deployed (*see* CUSTOMER EQUITY; CUSTOMER ANALYSIS) (Rust, Lemon, and Zeithaml, 2004). New and different applications of CLV analysis are constantly being developed (Fader, Hardie, and Lee, 2005; Jain and Singh, 2002).

There are some important caveats related to the use of CLV. At its best, it can be an accurate estimation of the future profitability of the customer examined but it is simply an estimate. CLV is highly dependent on the quality and accuracy of the data utilized. Depending on the complexity and sophistication of the CLV model and formulation used, the resulting estimate may not fully account for the very dynamic nature of the firm/customer relationship (Gupta *et al.*, 2006). In addition, except to the degree it is accurately reflected in the customer retention rate, CLV does not account for the action of competitors. Finally, since not all customers are the same, an aggregation bias can occur, especially when a generalized approach is used in the CLV estimation as there could be substantial differences in retention rates, servicing costs, and the amounts and rates of customer transactions.

Bibliography

Berger, P.D. and Nasr, N.I. (1998) Customer lifetime value: marketing models and applications. *Journal of Interactive Marketing*, **12** (1), 17–30.

Fader, P.S., Hardie, B.G.S., and Lee, K.L. (2005) RFM and CLV: using Iso-Value curves for customer base analysis. *Journal of Marketing Research*, **42**, 415–430.

Gupta, S., Hanssens, D., Hardie, B. *et al.* (2006) Modeling customer lifetime value. *Journal of Service Research*, **9**, 139–155.

Jain, D. and Singh, S.S. (2002) Customer lifetime value research in marketing: a review and future directions. *Journal of Interactive Marketing*, **16** (1), 34–46.

Rust, R.T., Lemon, K.N., and Zeithaml, V.A. (2004) Return on marketing: using customer equity to focus marketing strategy. *Journal of Marketing*, **68**, 109–127.

customer relationship management

V. Kumar

Introduction

A customer pulls into Harrah's Caesar's Palace in Las Vegas. A cheerful valet greets her by name and escorts her into the game room, circumventing the crowded lobby. Soon, a designated host approaches her and ensures that the service she is receiving matches her expectations. After her gaming session, she is escorted into the restaurant to her preferred table with her favorite wine waiting for her. This superior customer service is an instance of Harrah's building individual relationships with their high-value customers as a means to enjoy their continued patronage. Considering the satisfaction for customers and the subsequent repeat business it would generate for Harrah's, it is a win–win situation for both the company and the customers, and one that often leads to positive word of mouth and referrals for the company. This technique of developing a one-to-one relationship with customers is called *customer relationship management* (*CRM*). While the example above illustrates the customer experience (*see* THINKING DEEPER ABOUT CUSTOMER EXPERIENCE) in a business-to-consumer (B2C) setting, appropriate measures can be adopted to manage relationships with vendors in a business-to-business (B2B) setting.

What is CRM?

The above example illustrates the power of database marketing (*see* DATABASE MINING AND MARKETING) and how it can be harnessed to identify and analyze customer segments for valuable information that can be used to increase the impact of marketing campaigns for companies. Further, with advancements in technology, it is now possible to capture customer data *and* interact with the customers simultaneously, as demonstrated by Harrah's. Therefore, CRM extends the concept of database marketing to a customer level in order to develop profitable company-to-customer relationships. The concept of CRM can then be used to create strategies for interacting with each customer. Examples of such strategies include developing better relationships with profitable customers, locating and enticing new customers that will be profitable, and finding appropriate strategies for unprofitable customers, among others.

With this as the background, we can now see CRM from a business strategy perspective. By adopting such an approach, companies can gain long-term competitive advantage by optimally delivering value and satisfaction (*see* CUSTOMER SATISFACTION/DISSATISFACTION) to the customer and extract business value from the exchange. Kumar and Reinartz (2006) define CRM as follows:

> CRM is the strategic process of selecting the customers a firm can most profitably serve and of shaping the interactions between a company and these customers. The goal is to optimize the current and future value of the customers for the company.

The key components of this definition include the following:

- *Strategic process*: This means that CRM activities are initiated and managed starting from the top of the organization. CRM does not belong to any single department but needs contributions and reinforcements from all corporate functions. Furthermore, CRM is a continuing *process* that cannot be handled as just another software implementation project. It must be viewed as a continuous effort with the goal of the company becoming more customer focused.
- *Selection*: When the economic value of a customer is the basis for resource allocation, it is logical that firms focus on their most profitable or potentially profitable customers first. This is not about denying services to certain customers but about recognizing that there is a fit between a firm's offering and a segment's desires, behavior, and characteristics.
- *Interactions*: This means that the relationship between the customer and firm takes the form of an interactive dialogue. Information and goods are exchanged and, most importantly, the exchange evolves as a function of past exchanges. This is very different

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

from the view that firms sell one-off products and services to the customer (*see* SERVICES MARKETING STRATEGY).

- *Customers*: Depending on the industry and company, a customer can be an individual account, one or several segments within a market, or an entire market. Also, customers include not only end users but also intermediaries, like distributors, retailers, and so on.
- *Current and future value of the customer*: (*see* CUSTOMER SOLUTIONS) This means that firms are moving away from extracting their profit from single transactions to maximizing profits over a series of transactions. Thus, firms are starting to maximize CUSTOMER EQUITY – that is, the value of all their customer relationships to them. In this process, traditional measures such as market share are increasingly being replaced with new measures such as customer lifetime value (CLV).

CRM also involves automating and enhancing the customer-centric business processes of sales, marketing, and service. It not only deals with automating these processes but also focuses on ensuring that the front office applications improve CUSTOMER SATISFACTION, resulting in increased customer loyalty (*see* MARKETING METRICS) that directly affects the company's bottom line.

Loyalty Programs

Companies have long used loyalty programs in an effort to assess the current and future value of customers. A loyalty program is a marketing process that generates rewards to customers on the basis of their repeat purchases. Consumers who enter a loyalty program choose to focus more of their buying with a particular company, thereby forgoing the free choice they had otherwise. In exchange for concentrating their purchases with the company, they accumulate assets (e.g., "points") that are exchanged for products and services, typically associated with the focal firm.

All successful loyalty programs have a unified objective and stress on an important aspect in customer management (CM) – it is cheaper to market to existing customers than to acquire new ones. Considering the achievements of all loyalty programs thus far, we can summarize the key objectives of loyalty programs as follows:

- *Building true loyalty*: This refers to a greater commitment to the product or organization through the building of true loyalty, both attitudinal and behavioral. Attitudinal loyalty refers to the perceptions and attitudes that a customer has toward a particular product or service. Behavioral loyalty, on the other hand, refers to the observed action that the customers have demonstrated toward a particular product or service. This objective is also not easy to achieve because consumers can be fickle and economic benefits are important to them. Hence, building true loyalty is the goal of many loyalty programs.
- *Efficiency profits*: This refers to the comparison of profit consequences with and without loyalty programs to determine where the biggest gain is obtained in terms of return on investment (ROI) or profit. Therefore, the goal of cultivating profitable loyalty is an important objective.
- *Effectiveness profits*: These are long-term profit consequences realized through better knowledge of customer preferences over time. This goal ensures that profitable loyalty created is sustainable over the long term in generating value for customers.
- *Value alignment*: Finally, building true loyalty and cultivating profitable loyalty should align the cost of serving a particular customer with the value he or she brings to the firm. This ultimate goal enables a firm to serve its customers in the best manner possible.

However, one has to keep in mind that not all of these goals may be pursued by the loyalty programs. It is possible for loyalty programs to concentrate on certain goals alone. Further, companies have to measure and manage customer loyalty in order to develop longerlasting customer relationships.

Metrics to Measure and Manage Customer Loyalty

Traditionally, various metrics (see MARKETING METRICS) have been used to measure and

manage customer loyalty that are anticipated to be reasonable indicators of actual customer value. These metrics help firms prioritize their customers in a manner that allows a higher proportion of resources to be assigned to the customers who are expected to generate greater profits in the future. The following are some of the popular metrics used:

- *Recency, frequency, and monetary value (RFM)*: Recency refers to how long it has been since a customer last placed an order with the company, frequency refers to how often a customer orders from the company in a certain defined period and monetary value denotes the amount that a customer spends on an average transaction. The higher the computed score, the more profitable the customer is likely to be, in the future.
- *Past customer value (PCV)*: In this model, the value of a customer is determined on the basis of the total contribution (toward profits) provided by the customer in the past. Since products/services are bought at different points in time during the customer's lifetime, all transactions are adjusted for the time value of money.
- *Share of wallet (SOW)*: This refers to the proportion of category value accounted for by a focal brand or a focal firm within its base of buyers. It can be measured at the individual customer level or at an aggregate level.

The metrics discussed above score customers on the basis of their past purchase behavior and rank order customers from best to worst. However, research has shown that these traditional metrics are poorly correlated with future profitability. Therefore, we need a forward-looking metric that overcomes these limitations and helps firms accurately measure and manage customers profitably.

Need for a Forward-Looking Metric

We have found a solution to this problem in the CLV metric (*see* CUSTOMER LIFETIME VALUE (CLV)). Unlike other traditional measures that include only past contributions to profit, the merit of CLV rests on the fact that it is a forward-looking metric. The conceptual approach to compute CLV is illustrated in Figure 1.

The CLV metric can be computed using three main components: (i) contribution margin, (ii) marketing cost, and (iii) probability of purchase in a given time period (Reinartz and Kumar, 2003). Each of these models will have a set of drivers/predictors. These three models are estimated simultaneously. By applying this approach, managers can estimate the CLV for each customer in the firm. The calculation of CLV for all customers helps the firms to rank order customers on the basis of their contribution to the firm's profits. Such an analysis would help the firm to treat each customer differently on the basis of his or her contribution, rather than treating all the customers in a similar fashion.

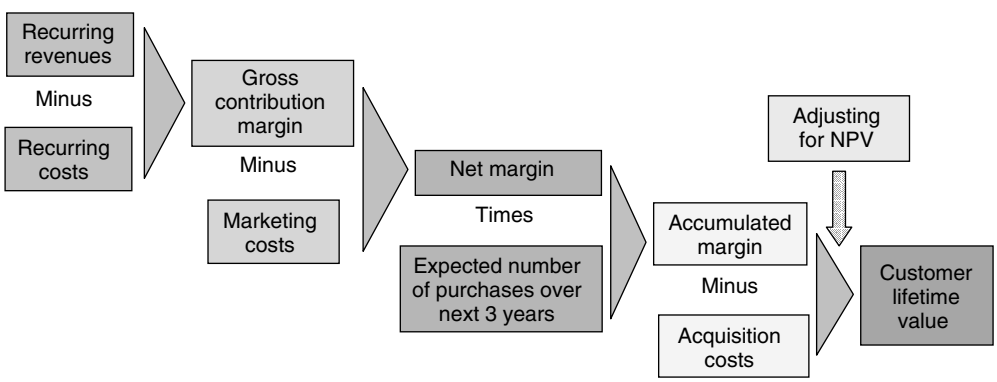


Figure 1 Approach to CLV measurement.

CLV assists marketers to adopt appropriate marketing activities today in order to increase future profitability. In addition, the computation can also be used to include prospects, not just current customers as used by the recency, frequency, and monetary value (RFM), share of wallet (SOW), and past customer value (PCV) metrics. Further, CLV is the only metric that incorporates into one, all the elements of revenue, expense, and customer behavior that drive profitability. This metric also manages to score over other metrics by adopting a customer-centric approach instead of a product-centric one, as the driver of profitability (Kumar, 2007).

One of the key outcomes in using CLV as a metric is in achieving convergence between marketing actions (e.g., contacts across various channels) and CRM. When managers have computed the CLV of their customers, they must look forward to maximize it in order to reap the benefits of the metric. The following section discusses the various strategies that firms can use to maximize CLV.

Strategies to Develop Profitable Customer Relationships

Using a powerful metric such as CLV, firms can address marketing issues with greater confidence. This metric will also help companies to develop important CRM strategies that will help them develop profitable relationships with their customers. Figure 2 illustrates the strategies, known as the *Wheel of Fortune* (Kumar, 2008).

Implementing these strategies is a cyclical process. The knowledge acquired in implementing these strategies should be used as the basis for deciding which customers to pursue in the future and how to select the most profitable set of customers. A brief synopsis of all the strategies is provided in this section.

Customer selection.

The strategy. The objective of customer selection is to detect and target customers/distributors on the basis of their value potential as opposed to other traditional metrics like RFM, PCV, and SOW. This strategy is applied when firms want to target individual customers or groups of customers. The reason for targeting these customers can be manifold, for example, for sending out a promotion or for inviting them to a special event. Finding the right targets for marketing resource allocation is at the heart of any CRM strategy. Smart targeting actions add value to the firm as well as to the customer.

Need for the strategy. First, the marketing budget of a firm is limited, and managers have to make choices as to where and on whom they should spend the limited resources available to them. Second, not all customers are equally profitable. An overwhelming share of profits is generated by a small percentage of customers. This necessitates targeting those customers with high profitability, and this is the basis of the customer selection strategy. Traditionally, firms rank order customers based on their profits and prioritize their resources based on this ranking. Among the metrics discussed in the previous section, the forward-looking CLV metric is the most successful in predicting future customer profits.

Benefits of this strategy. The performance of the traditional metrics (RFM, PCV, and SOW) versus the CLV metric in customer selection has been compared numerous times with CLV always offering higher levels of profitability. For example, Venkatesan and Kumar (2004) found that the net value generated by the customers who were selected based on the CLV score is about 45% greater than that generated through customers selected through other traditional metrics. This shows that using CLV to select customers is far more effective than using the traditional metrics. These findings provide substantial support for the usefulness of CLV as a metric for customer scoring and customer selection.

Managing loyalty and profitability simultaneously.

The strategy. Companies should manage customers who are *profitably* loyal and not just loyal. This is based on the finding by Reinartz and Kumar (2002) that not all loyal customers are necessarily profitable, and not all profitable customers are necessarily loyal. By managing loyalty and profitability simultaneously, this

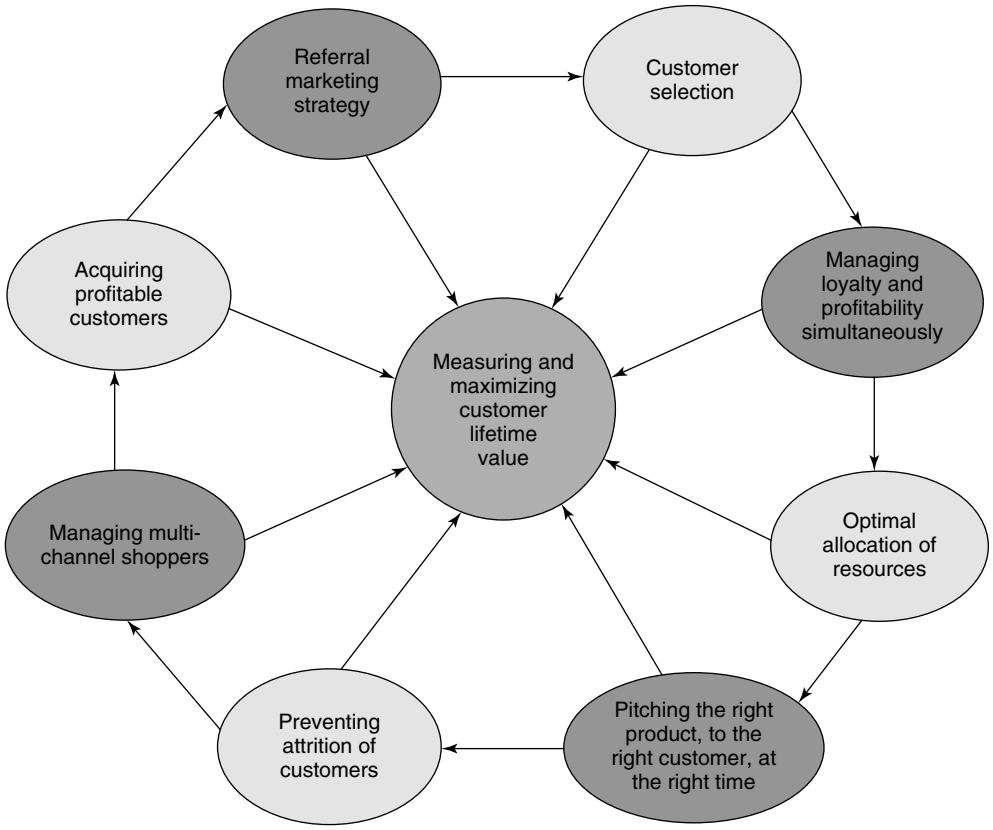


Figure 2 The Wheel of Fortune. *Source*: Adapted and updated from Kumar, V. (2008), *Managing Customers for Profits*, Upper Saddle River, NJ: Wharton School Publishing.

strategy clarifies the relationship between loyalty and profitability and provides managers with the required framework to manage loyalty and profitability at the same time. Implementation of this strategy calls for segmenting customers on the basis of their loyalty and profitability to the firm. There are several measures of loyalty and profitability that can be used for this purpose. Customers are segmented into cells on the basis of their loyalty and profitability levels. The four cells are as follows:

- *Cell 1 – Strangers (low profitable and short-term customers)*: This set of customers is not loyal to the firm and bring in little to no profits to the company. These customers have very little fit between the company's offerings, and have very little profit potential. They have to be identified very early on, and any investment toward building a relationship with them should be avoided. Profits should be made from every transaction with these customers.
- *Cell 2 – Butterflies (highly profitable and short-term customers)*: These are transient customers who can be very profitable, but often do not exhibit traditional behavioral loyalty. These customers tend to buy a lot in a short time period and then move on to new firms after making those transactions. These customers avoid building a long-term relationship with any single firm. Therefore, firms should adopt a more prudent approach in managing them–enjoy their profits while it lasts and stop investing in them at the right moment when they switch.

- *Cell 3 – True friends (high profitable and long-term customers)*: These customers are both loyal and profitable, and buy steadily and regularly (not too intensively) over time. True Friends are generally satisfied with existing arrangements with the firm, and are usually comfortable engaging with the firm's processes. Care should be taken in building relationships with these True Friends, since these customers have the highest potential to bring long-term profitability.
- *Cell 4 – Barnacles (low profitable and long-term customers)*: These customers, if managed unwisely, could prove to be a severe drain on the company's resources. The size and volume of the transactions made by these customers are too low to justify the cost incurred in marketing and maintaining their accounts. The primary step in managing these customers is to evaluate the size and share of their wallet. If the SOW is found to be low, specific up-selling and cross-selling can be done to extract profitability. However, if the size of wallet is small, then, strict cost control measures can reduce the loss for the firm.

Need for the strategy. This segmentation strategy is helpful for firms in building an effective loyalty program – one that is aimed at maximizing profitability for the firm. With the four segments as the base, firms can aim to (i) build and enhance behavioral loyalty by focusing on purchase behavior of individual customers and their respective contributions to firm profit, (ii) cultivate attitudinal loyalty by studying customers' attitudes toward the firm through customer feedback and focus groups, and (iii) link loyalty to profitability by using the CLV metric, and ensure a successful loyalty program. This would help companies to identify the high-value customers and spend the limited marketing resources on them.

Benefits of the strategy. All customers are not equally profitable. This should be remembered when putting a loyalty program in place. This strategy would help the firm reward customers depending on their level of profitability, such as a two-tiered loyalty program (Kumar and Shah, 2004). The tier-1 rewards are aimed at all customers based on their current and past purchase behavior. This tier-1 reward is a simple explicit way of rewarding customers and attracting new customers. Tier-2 rewards are aimed at influencing the future purchase behavior of customers. Tier-2 rewards are more selective and rewards customers to influence their behavioral and attitudinal loyalty.

Optimal allocation of resources.

The strategy. Ideally, firms should be investing only in customers who are profitable. However, many companies continue to spend resources on a large number of unprofitable customers. They either invest in customers who are easy to acquire but are not necessarily profitable or try to increase the retention rate of all their customers, thereby leading to wastage of limited resources.

Implementation of this strategy calls for identifying the most profitable customers, and the customers who are most responsive to marketing efforts. The second step is to determine the right mix of different channel contacts for each customer. This depends upon how responsive each customer is to these channels of communication (e-mail, direct mail, telephone, direct visit by a salesperson), and how cost effective these channels are. The next step in implementing this strategy is to decide how frequently the customer should be contacted, and what the intercontact time should be. Further, the various factors that affect customer behavior such as upgrading (to a higher product category), cross-selling (in a different category), and so on, need to be analyzed. Therefore, by carefully monitoring the purchase frequency of customers, the interpurchase time, and the contribution toward profits, managers can determine the frequency of marketing initiatives in order to maximize CLV.

Need for the strategy. Conventional wisdom says that a company's contact strategy and the frequency and modes of communications affect the customer's predicted purchase frequency. Further, it would help managers know the extent to which they should use the various contact channels. For instance, should there be an increase in face-to-face meetings

and a decrease in the frequency of direct mailers, or vice versa? Traditional marketing methods of flooding customers with product promotions will not only drain the company's limited marketing resources but also lead to customer estrangement. Venkatesan and Kumar (2004) have shown that different channels of communication have different costs associated with them and have to be employed at the optimum frequency and time intervals to maximize a customer's purchase frequency.

Benefits of the strategy. This strategy makes recommendations as to which customers to acquire and retain based on their predicted CLV. It also provides recommendations as to how to use the different channels of communication based on how responsive each customer is to these channels. And an optimal level of communication across the right mix of channels is recommended in order to achieve maximum profitability. When the channel reallocation was implemented in a B2B firm, the firm generated 100% more revenue and 70% more profit.

Pitching the right product to the right customer at the right time.

The strategy. Companies are constantly involved in predicting the next product that a customer is likely to buy. An ideal contact strategy is one where the firm is able to deliver a sales message that is relevant to the product that is likely to be purchased in the near future by a customer. This could be achieved by accurately predicting the purchase sequence.

Need for the strategy. Understanding the purchase sequence calls for analyzing past customer purchases and estimating the likelihood of future purchases to design optimal contact strategies. By analyzing and asking questions such as (i) in which product category the customer is likely to make a purchase, (ii) at what intervals and at what time period is the customer likely to make a purchase, and (iii) how much is the customer likely to spend or, in other words, how profitable the customer is likely to be, the strategy predicts the purchase sequence of each customer.

Benefits of the strategy. When Kumar, Venkatesan, and Reinartz (2006) tested this in a B2B setting, 85% of the customers predicted by this model to make a purchase actually went on to do so. In comparison, only 55% of the customers predicted by the traditional model actually made a purchase. When this strategy was implemented in the B2B setting, an increase in ROI of 160% was observed. Thus, this strategy suggests that efficient management of the purchase sequence not only increases revenue by accurately predicting and preempting a customer purchase but also minimizes cost by reducing the frequency of customer contacts.

Preventing customer attrition.

The strategy. As competition intensifies, companies find that their customers are very vulnerable and are likely to defect. Since customer churn has a tremendous impact on the firm's performance it is very critical for the firms to develop a successful strategy to prevent attrition. This strategy calls for calculating the probability of defection for each customer. With this as the base, and answers to questions such as: (i) Should we intervene? (ii) With which customers should we intervene? (iii) When should we intervene? (iv) Through which channel should we intervene? (v) What should we offer? companies can target those customers with intervention offers that would make them stay with the company for a longer period of time.

Need for the strategy. Customer attrition impacts a firm in several ways. First, firms experience a loss of revenue from the customers who have defected. Second, the opportunity cost for the firm to recover the acquisition cost on the defected customer is lost, which makes it difficult for the firm to break even. Third, the firm loses the opportunity to up-sell/cross-sell to customers who have defected, and this loss can be treated as a loss of potential revenue. Fourth, due to potential negative word of mouth, the firm loses out on future customers. This drains the firm's resources, which are already impacted by the loss of customers, mostly to competitors (*see* COMPETITOR ANALYSIS).

Benefits of the strategy. The effect of the customer intervention strategy was studied in a telecom firm using test and control groups. Using

this strategy, the firm realized a net revenue gain of $345 000, after accounting for the cost of intervention, and the ROI was close to 860%. Therefore, the key to retaining customers is to identify those who are likely to quit and intervene by sending appropriate messages.

Managing multichannel shoppers.
The strategy. This strategy demonstrates that multichannel shoppers (*see* MULTI-CHANNEL MARKETING) initiate more contacts with the firm, stay with the company for a longer duration, purchase more frequently, and are more receptive to contacts through multiple communication channels. Further, the strategy shows that there exists a nonlinear relationship between returns and multichannel shopping, and that there is a positive synergy toward multi-channel shopping when customers are contacted through various communication channels.

Need for the strategy. Companies are in a constant pursuit to grow and serve customers, and thereby use at least a few different channels. In many cases, these channels not only offer customers a chance to make purchases via multiple channels but they also offer customers the chance to search for product information in one or more channels and purchase in a completely different channel. Further, owing to advancements in logistics, increase in online selling formats, and diverse customer segments, firms are introducing multiple distribution channels to target the customer segments. Since each distribution channel services a different set of customers and provides varying levels of services, this approach can lead to a reduction in the overall service cost, resulting in an increase in profitability for the firm. Therefore, it would be profitable for firms to start operating across multiple channels and thereby target the multichannel shoppers.

Benefits of the strategy. Using this strategy, Kumar and Venkatesan (2005) showed that as a customer shops across more channels (from one channel to four channels), the customer (i) spends more revenue with the firm, (ii) spends a higher proportion on the focal firm (rather than with a competitor), (iii) has a higher past profitability (which is correlated with future profitability), and (iv) has a higher likelihood of buying in the future. By adopting the test and control group experiments, it was found that the net gain in revenue for the firm due to the addition of one more channels was on average about 80%. After accounting for the costs, the return on investment was about eight times or 800%. Therefore, when firms contact the right customers, at the right time, to encourage adopting another channel, it results in higher profitability.

Acquiring profitable customers.
The strategy. The main goal of this strategy is to acquire profitable customers, make them stay longer with the company, and derive more profits from them over their lifetime with the company. Using CLV, this strategy optimizes the acquisition/retention costs of customers and links such efforts to overall profitability. Such an exercise helps firms to decide which customers are worth acquiring/retaining and which dormant customers should be pursued to come back to the firm. Firms should use customer profiles to identify the customers who are most likely to be profitable. This can be achieved by identifying customers with similar characteristics as existing high-CLV customers, and by adopting an appropriate marketing strategy.

Need for this strategy. The need for this strategy arises from the fundamental question to which every company seeks answers – whom to acquire and retain? While making direct marketing investment decisions, many marketers still overemphasize short-term cost over long-term gain. This leads to a situation where companies would pursue customers who are cheap to acquire and cheap to retain without essentially being profitable. Conventionally, most companies use customer acquisition and retention rates to measure their marketing performance. This approach could diminish returns to the firms as they might be spending more on acquiring and retaining a customer, than on what the customer brings in as revenue. Further, different groups of customers require different levels of acquisition/retention spending to maintain their relationship. Therefore, managers need a strategy that strikes a balance by efficiently allocating resources

between customer acquisition and customer retention.

Benefits of this strategy. By observing the customer behavior of a catalog retailer, Thomas, Reinartz, and Kumar (2004) segmented customers on the basis of their cost of acquisition and cost of retention. They found that the largest segment (32%) was made up of customers who were easy to acquire and retain. But, they accounted for only 20% of the total profits. On the other hand, 40% of the total profits came from the smallest group of customers (15%) who were expensive to acquire but cheap to retain. Therefore, linking acquisition and retention to profitability helps the firm to target and retain profitable prospects and customers. This demonstrates that optimal allocation of resources helps companies increase their revenue, balance acquisition and retention strategies, and manage customers profitably.

Referral marketing strategy.

The strategy. This strategy introduces the use of the Customer Referral Value (CRV) metric, which is defined as the value of the referral behavior for a specific customer, as the basis for managing and maximizing customer referral behavior. This metric advises managers that customers be valued on the basis of their indirect impact on the firm's profits, through savings in acquisition costs and addition of new customers by way of customer referral.

Need for the strategy. Past research studies have suggested the existence of a link between a customer's willingness to recommend and the growth in firm's profits (Reichheld, 2003). Specifically, the ultimate question for a firm is – would you recommend us? Relying only on this question leads to two problems: (i) this question would highlight only the correlation and not the causation of referrals and increase in profits, and (ii) this question would identify the referral behavior on an aggregate level and not on an individual customer basis. Recently, Kumar, Petersen, and Leone (2007) conducted a field study to better understand the determinants of customer referral behavior and its impact on firm profits by computing CRV.

Benefits of the strategy. It is important to note that while CLV indicates the actual value of purchases provided by a customer to the firm, CRV computes the influence that the customer has on other potential customers. Therefore, firms should measure and manage CLV and CRV simultaneously to drive more value for the firm. To develop an effective CRM strategy, a telecommunications firm computed the CLV and CRV for a sample of 9900 customers and segmented them into four cells of a 2×2 matrix. The four cells are (i) affluents (high CLV and a low CRV), (ii) misers (low CLV and a low CRV), (iii) advocates (low CLV and a high CRV), and (iv) champions (high CLV and a low CRV).

Therefore, the managerial implications in maximizing CLV and CRV simultaneously is that (i) not all behaviorally loyal customers recommend new customers, (ii) not all customers who recommend new customers are behaviorally loyal, and (iii) focusing only on high-CLV customers for increasing the number of referrals is not an optimal strategy. The key to success is to encourage customers to build social networks and to trigger them, along with customers who already have strong social networks, to talk to individuals in their network by using marketing programs that reward word of mouth and/or referrals.

The strategies discussed here are proven strategies that can help companies in building profitable longerlasting customer relationships. Even though these strategies can yield great results when implemented separately, they work best when implemented as a whole. The knowledge acquired from implementing each strategy can be used to improve on other strategies, and this information can be collectively used to refine future CRM initiatives.

Future of Customer Relationship Management

Until now, CRM strategies adopted by companies have always focused on cost savings in managing customers. Increasingly, companies are beginning to realize the potential of cultivating profitable customer loyalty that can help them increase revenue and profit. When implementing CRM strategies, firms have to adopt a customer-focused view that considers

their needs, lifetime value, and profitability. In this regard, a product-focused view where decisions are made with the product in reference would be of little help to the companies. Such a managerial transformation requires indentifying necessary components and interfaces for delivery of products/services to the customer, human resources to implement and manage the process, and appropriate metrics that can measure and maximize the impact on any change in the firm's consumer base. The future of CRM, therefore, lies in the ability of firms to accurately assess customer feedback and market trends through an interactive marketing (*see* DIRECT AND INTERACTIVE MARKETING) approach and the capacity to respond instantly to changing market and customer needs.

Acknowledgments

The author wishes to thank Bharath Rajan in the preparation of this manuscript. The author thanks several financial services, pharmaceutical, grocery, high tech, telecommunication, and retailer firms for agreeing to share their customer databases for this study.

Bibliography

Kumar, V. (2007) Customer lifetime value: the path to profitability. *Foundations and Trends in Marketing*, **2** (1), 1–95.

Kumar, V. (2008) *Managing Customers for Profits*, Wharton School Publishing, Upper Saddle River, NJ.

Kumar, V., Petersen, J.A., and Leone, R. (2007) How valuable is word of mouth? *Harvard Business Review*, **85** (10), 139–146.

Kumar, V. and Reinartz, W.J. (2006) *Customer Relationship Management: A Databased Approach*, 1st edn, John Wiley & Sons, Inc., New York.

Kumar, V. and Shah, D. (2004) Building and sustaining profitable customer loyalty for the 21st Century. *Journal of Retailing*, **80** (4), 317–330.

Kumar, V., Shah, D., and Venkatesan, R. (2006) Managing retailer profitability: one customer at a time! *Journal of Retailing*, **82** (4), 277–294.

Kumar, V. and Venkatesan, R. (2005) Who are the multi-channel shoppers and how do they perform?: correlates of multi-channel shopping behavior. *Journal of Interactive Marketing*, **19**, 44–62.

Kumar, V., Venkatesan, R., and Reinartz, W.J. (2006) Knowing what to sell, when, and to whom. *Harvard Business Review*, **84** (3), 131–137.

Reichheld, F. (2003) The one number you need to grow. *Harvard Business Review*, **81** (12), 46–54.

Reinartz, W.J. and Kumar, V. (2002) The mismanagement of customer loyalty. *Harvard Business Review*, **80** (7), 86.

Reinartz, W.J. and Kumar, V. (2003) The impact of customer relationship characteristics on profitable lifetime duration. *Journal of Marketing*, **67** (1), 77–99.

Thomas, J., Reinartz, W.J., and Kumar, V. (2004) Getting the most out of all your customers. *Harvard Business Review*, **82** (7/8), 116–123.

Venkatesan, R. and Kumar, V. (2004) A customer lifetime value framework for customer selections and resource allocation strategy. *Journal of Marketing*, **68** (4), 106–125.

customer satisfaction/dissatisfaction

Steven P. Brown and William Zahn

CUSTOMER SATISFACTION

Customer satisfaction is a tacit or explicit assessment by customers that they have received positive outcomes and fair value for sacrifices made in acquiring and utilizing a product or service. In fundamental respects, it lies at the heart of marketing thought and practice. Given its centrality to marketing thought and practice, it is not surprising that customer satisfaction has been considered from various perspectives and at different levels of analysis in marketing research. Customer satisfaction research accounts for a substantial proportion of total revenues of leading market research firms (Oliver, 1999), and academic research on the topic provides an integrating theme that can be traced across research paradigms and substantive fields of interest.

Customer satisfaction might fairly be said to constitute the conceptual centerpiece of marketing thought and practice, or at least to have as substantial a claim on that distinction as any other construct. In this article, we organize the many viewpoints that have been taken on customer satisfaction into three categories: philosophical, psychological, and managerial.

The philosophical perspective. In his classic article on "Marketing Myopia," Levitt (1960) argues that the brilliance for which we remember Henry Ford should *not* be for having revolutionized production processes; it should be for his marketing prowess. Citing Ford's own writings, Levitt notes that Ford first came to understand that he could sell millions of cars priced at $500, and only then set out to devise a way to profitably satisfy this latent demand. Assembly line production became the means to this end, but the driving innovation was that the goal of satisfying customer wants and needs at a profit at a price they were willing to pay dictated product design and the means of production. While he was apparently denying it (by saying "customers can have it in any color they want, as long as it is black"), Ford had become the father of MARKET ORIENTATION.

Levitt's (1960) article is frequently cited as the seminal statement that customer satisfaction and customer orientation lie at the heart of business and constitute its essential purpose. This philosophical position has since been thoroughly interwoven into marketing thought (if not always into practice). It is reflected in Levitt's observations that "The view that an industry is a customer-satisfying process, not a goods producing process, is vital for all businesspeople to understand" (p. 55), and that "...the entire corporation must be viewed as a customer-creating and customer-satisfying organism. Management must think of itself not as producing products but as providing customer-creating value satisfactions" (p. 56; *see* A FRAMEWORK FOR CREATING VALUE PROPOSITIONS).

Levitt's philosophical perspective has been updated by Kohli and Jaworski (1990); Narver and Slater (1990), and others in their descriptions of market-oriented organizations (*see* MARKET ORIENTATION). Their contributions, which we discuss in more detail in the section Managerial Perspectives, have supported the philosophical perspective and sustained the role of customer satisfaction as the centerpiece of marketing thought and practice.

Noting that satisfied customers often do not remain brand loyal (*see* BRAND EQUITY; BRAND GROWTH STRATEGY), some writers (e.g., Reicheld, 1996) have urged scholars and managers to redirect their attention from customer satisfaction to loyalty. Addressing this issue, Oliver (1999) underscores the indispensable role of customer satisfaction in the development and maintenance of loyalty (*see* CUSTOMER EQUITY). He likens the metamorphosis from satisfaction into loyalty to the germination of a seed into a flower when nurturing elements (e.g., differentiated quality, social support, and consumer resistance to competitive offerings) are present. He further notes that in many product categories (e.g., those in which real differentiation, customer identification, and social support are not feasible) true loyalty is not realistically achievable, and satisfaction should remain the firm's primary objective.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

The psychological perspective. Feelings and judgments of satisfaction (e.g., with life, work, or relationships, as well as with products and services) importantly influence the quality and meaning of life experience (*see* THINKING DEEPER ABOUT CUSTOMER EXPERIENCE). As a result, considerable research effort has endeavored to understand the processes that determine satisfaction judgments.

From the psychological perspective, Oliver (1999) defines customer satisfaction as "pleasurable fulfillment." Research on processes resulting in customer satisfaction has focused largely on the disconfirmation of expectations paradigm. In this view, customers are believed to have preexisting expectations about the performance of products and services and to derive satisfaction judgments based on how their actual consumption experience measures up against these expectations. Positive disconfirmation produces satisfaction, whereas negative disconfirmation leads to dissatisfaction. Although the preponderance of research has shared this general premise, different researchers have posited and tested different types of standards and expectations against which to compare product performance (e.g., desires, ideals, experience-based norms, comparison levels of alternatives).

The disconfirmation of expectations paradigm has been criticized for being too focused on single consumption experiences, excessively cognitive, narrowly individualistic, and overly abstracted from consumption context. Seeking to broaden the scope of customer satisfaction research, Fournier and Mick (1999) contextualized the psychological perspective using ethnographic methods. Although they found some support for disconfirmation processes, their investigation also led to conclusions that customer satisfaction is partially rooted in the life situation and social context of the consumer (*see* THINKING DEEPER ABOUT CUSTOMER EXPERIENCE). Products that facilitate progress toward higher order goals (e.g., being a good parent, being a consummate professional) and enrich the enjoyment and meaning of social interactions produce satisfaction. Also, products that inspire awe, produce positive surprise and delight, eliminate expected negative consequences of consumption, or become integral to the consumer's existence deliver high levels of satisfaction. Fournier and Mick (1999) also found customer satisfaction to result from dynamic processes that are shaped through extended interactions with products over time. In all, this research usefully broadened the psychological perspective on customer satisfaction by highlighting the limitations of the dominant disconfirmation paradigm (while not totally rejecting it) and emphasizing the goal-directed and socioemotional aspects of consumption. They also highlighted what they perceived as an integral linkage between customer satisfaction and life satisfaction (particularly in the sense that products that further life goals produce customer satisfaction).

The managerial perspective. The managerial perspective on customer satisfaction is concerned with the practical realization of the philosophical perspective. If from a philosophical perspective customer satisfaction is a primary goal that shapes product development and means of production, then the managerial perspective is concerned with the means by which this orientation can be implemented most effectively. The ubiquity of customer satisfaction measurement in industry suggests the extent to which executives accept the philosophical perspective and strive to implement it. In contrast to research on the psychological perspective, which is typically conducted at the individual level of analysis, research on the managerial perspective aggregates customer satisfaction scores to the organizational level and seeks to identify antecedents and outcomes of satisfaction at this level.

Fornell and Wernerfelt (1987) provided an important impetus to the managerial perspective on customer satisfaction by indicating the substantial returns firms can realize from "defensive marketing" (i.e., customer retention resulting from high levels of satisfaction; *see* CUSTOMER RELATIONSHIP MANAGEMENT). Reicheld (1996) provided a further impetus with an analysis indicating that an improvement of 5% in customer retention can result in profitability improvements ranging from 25 to 95% (depending on industry; *see* CUSTOMER LIFETIME VALUE (CLV)). Although customer loyalty (rather than satisfaction per se) is the

focal variable in these analyses, as previously noted, customer satisfaction is the sine qua non underlying loyalty (in the sense of attitudinal commitment to the brand).

Development of the American Customer Satisfaction Index (ACSI) at the University of Michigan, which benchmarks over 200 firms based on customer satisfaction scores, has importantly advanced research from the management perspective. Studies based largely on the ACSI database have established satisfaction as a leading indicator of financial performance metrics, such as operating margin, return on investment, and shareholder value (e.g., Anderson, Fornell and Mazvancheryl, 2004; *see* MARKETING METRICS). Most of this research has shown meaningful variation in the strength of relationship between satisfaction and financial performance across industries. Other research has shown that customer satisfaction also improves the efficiency of future advertising and promotion investments (Luo and Homburg, 2007; *see* COMMUNICATIONS BUDGETING).

Conceptual and empirical research on market orientation (e.g., Kohli and Jaworski, 1990; Narver and Slater, 1990) has extended Levitt's work by identifying processes that characterize market-oriented businesses and relating them to organizational performance. In Kohli and Jaworski's (1990) framework, these processes include intelligence generation, dissemination across functions, and coordinated response to market developments. Parsing the conceptual domain of market orientation somewhat differently, Narver and Slater's framework includes customer orientation, competitor orientation, and interfunctional coordination. Both of these views of market orientation have been shown to be positively related to firm profitability.

At a finer grained level of analysis, organizations frequently seek to leverage their investments in satisfaction research by identifying specific process improvements that will most effectively improve customer satisfaction. Product attributes that correlate highly with customer satisfaction and on which performance has considerable room for improvement provide the best targets.

Survey-based efforts to improve customer satisfaction have been criticized for being backward looking, as customer feedback reflects past actions and may provide no hint of impending discontinuous innovations that could render present markets obsolete. In this regard, it is important for firms to maintain a broad vision in their collection and dissemination of market information and not rely excessively on customer feedback about current product forms. Such feedback must be interpreted in the context of latent needs that customers may not be able to articulate easily, and technological advances that could result in new and radically changed product forms that serve the same needs (*see* INNOVATION DIFFUSION). This point was central to Levitt's (1960) original statement of the marketing concept.

Summary

As the core of the marketing concept, customer satisfaction has a reasonable claim to being the conceptual centerpiece of marketing thought and practice. Although some have reasonably argued that "satisfaction is not enough" (i.e., many satisfied customers switch to competitors' products), deep brand loyalty begins with the experience of satisfaction, and without it true loyalty cannot develop. Conversely, dissatisfaction leads to customer defections, negative word of mouth, and even customer efforts to harm the brand.

Given the centrality of customer satisfaction to marketing thought and practice, it is not surprising that it has inspired a great abundance and diversity of research. As our review suggests, widespread acceptance of the marketing concept reflects a philosophical perspective in which delivering customer satisfaction is the fundamental purpose of business. This perspective suggests the importance of understanding the psychological processes that lead to customer satisfaction. In this regard, we conclude (in agreement with Fournier and Mick, 1999) that the dominance of the expectancy disconfirmation paradigm has restricted progress toward a broader and more contextualized appreciation of socioemotional factors that shape the experience of satisfaction. Finally, research conducted at the organizational level of analysis has shown that customer satisfaction is related to a range of financial performance metrics. These findings support the philosophical perspective as

originally articulated by Levitt (1960). As the psychological perspective has suggested, satisfaction importantly influences the nature and quality of people's life experiences, and delivering it reliably to customers is fundamental to firms' financial performance.

Bibliography

Anderson, E.W., Fornell, C. and Mazvancheryl, S.K. (2004) Customer satisfaction and shareholder value. *Journal of Marketing*, **68**, 172–185.

Fornell, C. and Wernerfelt, B. (1987) Defensive marketing strategy by customer complaint management. *Journal of Marketing Research*, **24**, 337–346.

Fournier, S. and Mick, D.G. (1999) Rediscovering satisfaction. *Journal of Marketing*, **63**, 5–23.

Kohli, A.K. and Jaworski, B.J. (1990) Market orientation: the construct, research propositions, and managerial implications. *Journal of Marketing*, **54** (2), 1–18.

Levitt, T. (1960) Marketing myopia. *Harvard Business Review*, **38**, 45–56.

Luo, X. and Homburg, C. (2007) Neglected outcomes of customer satisfaction. *Journal of Marketing*, **71**, 133–149.

Narver, J.C. and Slater, S.F. (1990) The effect of a market orientation on business profitability. *Journal of Marketing*, **54** (4), 20–35.

Oliver, R.L. (1999) Whence customer loyalty. *Journal of Marketing*, **63** (Special Issue), 33–44.

Reicheld, F. (1996) *The Loyalty Effect*, Harvard Business School Press, Boston.

customer solutions

Sundar G. Bharadwaj

In the last decade, firms have increasingly moved in the direction of offering solutions to their customers. Recent actions by Dell to purchase Perot Systems, HP in purchasing EDS, or Xerox's purchase of Affiliated Computer Systems were all motivated by the need to offer solutions. This strategic action which initially began in the technology sector has broadened out to other sectors including financial services, travel, healthcare, automotive, and others. Actions such as using reverse auctions, group purchasing, global sourcing, and supplier tiering have increasingly put pressure on supplier firm prices and profitability. To overcome the perceived commoditization that such purchasing approaches impose on their offerings, supplier firms have added solutions to their portfolio with the explicit hope of differentiating their offerings and appropriating the value created through better prices. However, in a vast majority of instances, the expected premium prices have not materialized. The fundamental problem has been the inherent confusion that supplier firms have faced with regard to what a solution really is. Customers have tended to perceive the solutions offerings of their supplier firms merely as bundles largely consisting of products and service that are at best customized. Consequently, customers have not tended to provide the premiums that the suppliers have anticipated for such offerings, while in fact they have demanded a discount for buying a bundle. On the other hand, suppliers have faced increased costs in designing customized bundles and with lower prices their bottomlines have been significantly hurt. The author has drawn upon a study on both customers and managers from supplier firms to present a new way of defining and delivering solutions (Tuli, Kohli, and Bharadwaj, 2007).

The research indicates that while most supplier firms think of a solution as a customized and integrated bundle of goods and services, most customers, in sharp contrast, think of a solution as *a set of relational processes comprising (i) customer requirements definition, (ii) customization and integration of goods and/or services, (iii) their deployment, and (iv) postdeployment customer support, all aimed at meeting a customer's business needs*. The contrasting viewpoint leads suppliers to under-emphasize customer requirements definition and postdeployment support that are crucial to many customers. Suppliers' inattention to these relational processes appears to result in ineffective solutions, dissatisfied customers, and lower profitability. Viewing a solution as a set of four relational processes can help suppliers organize themselves to deliver greater value to customers, determine their cost-to-serve more accurately, and appropriate the value through more favorable prices.

Understanding a customer's business and needs is a nontrivial task. Firms need to translate the benefits of the solution in terms of outcomes that match the customer needs closely. For example, Lorenzo Zambrano, the CEO of Cemex, a cement manufacturer, realized that their real business was helping customers complete their construction project, not buy cement. While their goal was to sell more cement to low-income customers in Mexico, they also realized (after conducting participant observation research by living with them for a year in Mexico) that customers relied on repatriated funds from family members in the United States. Repatriation was a costly process and transaction costs were as much as 12%. Moreover, the families in Mexico had little expertise overseeing construction process. Cemex innovated a solution for such customers, wherein they provided construction advice, repatriated funds directly to contractors (lowering transactions costs substantially), and developed precast cement blocks and a new type of low-cost slurry to meet the needs of the low-income target market. An obvious managerial lesson is that to encourage an outcome-based thinking in the organization, firms need to get their employees to think of a whole range of related needs a customer faces in completing a task.

The effectiveness of a solution depends on both supplier and customer related structural and process oriented actions as well as characteristics. In particular, suppliers can enhance their ability to provide solutions by (i) adopting contingent hierarchies, in which a business unit with the most to contribute to a particular

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

solution has the authority over other business units for providing that solution; (ii) prespecifying the organizational process for developing solutions; (iii) aligning incentives of functions and units around providing effective solutions; (iv) emphasizing the documentation of customer engagements; and (v) limiting customer contact personnel turnover.

The effectiveness of a solution also depends on a customer's willingness to (i) adapt to a supplier's offerings, (ii) share information about its internal operations with a supplier, and (iii) share relevant internal political considerations with a supplier. In summary, these factors serve as the segmentation and targeting guideline for suppliers.

Bibliography

Tuli, K., Kohli, A. and Bharadwaj, S. (2007) Rethinking customer solutions: from product bundles to relational processes. *Journal of Marketing*, **71**, 1–17.

database mining and marketing

Maytal Saar-Tsechansky

Database mining refers to the data-driven process of extracting useful patterns or relationships from large databases, usually by means of sophistical statistical techniques. In the context of marketing, database mining and marketing refers to the application of data mining techniques with the objective of informing marketing decisions, such as direct targeting decisions. In contrast to some traditional marketing modeling approaches, data mining is principally data-driven. That is, it discovers and unravels patterns in data and is not guided by a theory of the underlying domain.

While most data mining techniques are generic and were not specifically developed for marketing applications, marketing has proven to be a popular application for database mining. Indeed, applications of data mining in marketing have been highly successful, and some have even become standard industry solutions. The reasons for this success include (i) the availability of large amounts of clean and structured data (such as scanner, purchase, and click-through data), (ii) open-mindedness regarding the interpretability of models (black box approaches are often acceptable), (iii) moderate requirements regarding the performance of data mining (i.e., the cost of decision or prediction errors is typically not very high), and (iv) significant, measurable, and immediate monetary benefits often result from the application.

Most marketing applications of data mining techniques utilize customer-specific demographic information, customer behavioral records, or a combination of both. The adoption of the Internet as a sales platform and a virtual environment offers a plethora of additional data sources including user-generated content including reviews, blogs, and Facebook, and social networks, all of which bear significant relevance for marketing problems. Below are illustrative marketing database mining applications that have been adopted widely in industry.

Market basket analysis: This employs algorithms that efficiently extract common co-occurrences of product purchases. More formally, market basket analysis employs algorithms that extract association rules of the form: if A then B, where A and B are sets of products. Such rules are often interpreted as: if one buys the products in A, then one is more likely (than a customer drawn at random) to buy the items in B. Market basket analysis has been applied to support decisions in store design, promotion, and even product recommendations.

Propensity models: These are predictive models that are typically applied to estimate the likelihood of a certain event, such as the likelihood that a customer will purchase a given product, or discontinue a service. Modeling techniques that have performed well include support vector machines, which are particularly effective for large sets of independent variables, and ensemble modeling approaches such as bagging, which effectively exploit extensive transactional histories. Propensity models are useful tools for designing mailing lists and optimizing the use of sales resources by prioritizing customers for targeting.

Customer lifetime value models: These go beyond modeling a customer's single purchase decision and address the long-term effect of retention efforts. Insights from customer lifetime value modeling can be used to optimize pricing and quality of service decisions. *Wallet models* serve a similar purpose and estimate the total budget that a customer has at his or her disposal for a given product category. This information is of general strategic value and can affect both product- and customer-specific pricing and marketing actions.

Ad placement: This is a more recent marketing application of data mining. The objective of ad placement is to match an ad with a relevant online content so as to increase the likelihood that a customer will be influenced by the ad. In this application, methods from information retrieval analysis are often used both for representing content and ads using keyword vectors and for

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

capturing the "proximity" of a given content to a set of keywords representing a given ad.

Viral marketing: This capitalizes on customer social networks that are either associated with a product (such as social networks derived by telecommunication providers based on customers' interactions) or social networks derived from public data sources, such as blogs. Data mining has demonstrated that explicit social networks as well as networks derived not from actual interactions but based on behavioral similarities provide valuable cues to help predict consumers' preferences. For example, network neighbors of a customer who has purchased a service were found to be more likely to order the same service in the future. Such networks have also been found to be effective for identifying targets for brand advertisements.

demand elasticity

Ashutosh Prasad

Introduction

Pricing is an important component of the MARKETING MIX. Here we cover a well established concept with direct implications for PRICING STRATEGY, namely, the price elasticity of demand. Also discussed are the related concepts of cross-price elasticity and income elasticity of demand.

The price elasticity of demand is a measure of consumers' sensitivity to price changes. Demand is said to be more (respectively, less) elastic if there is a large (respectively, small) change in quantity demanded in response to a price change.

The concept of elasticity was introduced by Marshall in 1882. Bartels (1976, p. 11) in *The History of Marketing Thought* writes that "The concept of elasticity was the result of a long development in economic thought. Adam Smith had recognized that demand was not a fixed amount at all prices. Malthus, on the other hand, had a better appreciation of some of the subjective character of demand and recognized what he called "intensity" of demand. Jevons and the Austrians expanded that idea into a theory of diminishing utility. Alfred Marshall later generalized it in the concept of elasticity. His interpretation of demand has long been used by marketing writers as a theoretical basis for selling, advertising, and the promotional work of marketing in general."

Definition of elasticity. The price elasticity of demand (henceforth *elasticity*) is defined as

$$\varepsilon(P) = \frac{\partial D(P)/D(P)}{\partial P/P} = \frac{P}{D(P)} \times \frac{\partial D(P)}{\partial P} \qquad (1)$$

where,

P is the price of the product;

$D(P)$ is the quantity demanded at price P, that is, the demand curve for the product. Function arguments are suppressed hereafter when no confusion arises. The possible dependence of demand on additional variables, such as prices of other goods and customers' incomes, explains the use of partial derivatives.

To apply Equation 1 we need a differentiable demand function which can be obtained either from estimation or by assumption. If the demand curve is unknown but a few discrete points on it are given, then an approximation of elasticity, sometimes called the *arc* elasticity of demand, is useful:

$$\varepsilon \approx \frac{\Delta D/D}{\Delta P/P} = \frac{\%\text{ change in demand}}{\%\text{ change in price}} \qquad (2)$$

In this context the definition in Equation 1 is referred to as the *point elasticity of demand*. A derivation of the point elasticity and arc elasticity is provided in Example 1.

Example 1. A store observes that the quantity demanded for a product is D_1 units when it is priced at P_1 and D_2 units when it is priced at P_2. What is the elasticity?

We can compute the point elasticity by assuming a two-parameter demand function that passes through the data points, followed by an application of the elasticity formula in Equation 1. The answer will depend on the chosen demand function. Thus, for the demand function $D(P) = a - bP$, it is $\varepsilon(P) = \frac{(D_1-D_2)/D(P)}{(P_1-P_2)/P}$. Note that the elasticity varies along the demand curve. For example, at the midpoint of the data, $(D_1 + D_2)/2$ and $(P_1 + P_2)/2$, the elasticity is $\varepsilon = \frac{D_1-D_2}{D_1+D_2} \times \frac{P_1+P_2}{P_1-P_2}$.

The alternative to assuming a functional form for demand is to compute the arc elasticity. When using the arc elasticity formula, often the data midpoint $\overline{D} = (D_1 + D_2)/2$ and $\overline{P} = (P_1 + P_2)/2$ is chosen as the reference point from which percentage changes are computed. Then, applying the arc elasticity formula yields $\varepsilon = \frac{D_1-D_2}{D_1+D_2} \times \frac{P_1+P_2}{P_1-P_2}$. This illustrates that when evaluated at the data midpoint, arc elasticity and point elasticity with linear demand are equivalent. □

If there are more than two data points, regression analysis can be used to first fit a demand curve to the data and then elasticity can be obtained.

Elastic and inelastic demand. From its definition, elasticity will be negative for a downward sloping demand curve. Since this is usually the case, the minus sign is often ignored. For example, when we say that demand is "more

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Estimated price elasticities of demand.

Inelastic Demand	*Approximately Unit Elastic*	*Elastic*
Gasoline and oil, ns	Motion pictures, −0.87	Foreign travel (lr), −1.77
Water consumption, −0.20	Medical care (lr), −0.91	Tobacco products (lr), −1.89
Medical care, −0.31	Shoes and footwear, −0.91	Restaurant meals, −2.27
Legal services, −0.40	Laundering services, −0.93	
Jewelry and watches, −0.41		
Tobacco products, −0.46		
Funeral and Burial expenses, −0.47		
Taxi, −0.63		

(ns = not significant; lr = long-run)

elastic" or has "higher elasticity" the reference is to the magnitude of elasticity. The magnitude, or absolute value, of elasticity is denoted by $|\varepsilon|$. Demand is said to be *elastic* if $|\varepsilon| > 1$, *inelastic* if $|\varepsilon| < 1$, and *unit elastic* if $|\varepsilon| = 1$. Furthermore, demand is *perfectly elastic* if $|\varepsilon| = \infty$ and *perfectly inelastic* if $\varepsilon = 0$. An example of the latter occurs when the demand curve is parallel to the price axis.

Table 1 was constructed from results selected from Houthakker and Taylor's (1970) analysis of US consumption data to show a few product categories with elastic, inelastic, and unit elastic demands.

A demand curve will generally exhibit different elasticities at different price points. This is illustrated by Example 2.

Example 2. What is the price elasticity of the linear demand function, $D = A - BP$, where A and B are positive constants?

The formula for elasticity yields $\varepsilon = \frac{-BP}{A-BP}$. Thus, elasticity is a function of the price. Figure 1 shows that the linear demand curve is divided into elastic and inelastic regions.

At its highest value, demand has zero elasticity and at its lowest value it is perfectly elastic. As a curiosity, the elasticity at any point on the linear demand curve is equal to the length of the line segment above that point divided by the length below it. □

Measuring demand sensitivity. As an alternative to computing the elasticity, a related measure of consumers' sensitivity to price changes is the slope of the demand curve $\partial D/\partial P$. Of these two measures, elasticity is often preferred because it is a unitless quantity, which facilitates comparisons across different products, currencies, and quantities. However, unlike the slope, it is not straightforward to assess elasticity by eyeballing the demand curve, with the exception that when the demand curve is drawn on a log–log graph, the slope will be exactly the elasticity. To see this, consider the function $y = f(x)$, where $y = \ln D$ and $x = \ln P$ are the transformed demand and price variables. Then the slope of this function is given by

$$\frac{\partial y}{\partial x} = \frac{\partial \ln D}{\partial \ln P} = \frac{\partial D/D}{\partial P/P} = \varepsilon$$

One can derive the demand function requirements for elasticity to be constant. Assuming that elasticity is constant, the above equation $\varepsilon = \partial \ln D/\partial \ln P$ can be solved as a differential

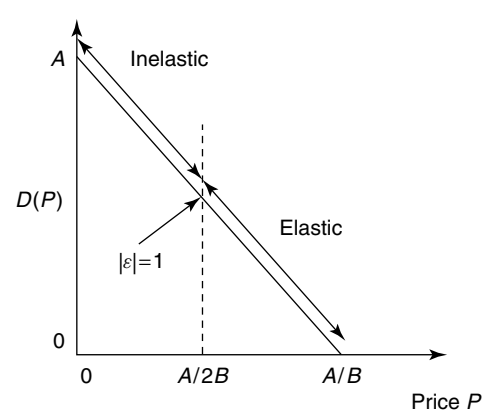


Figure 1 Elastic and inelastic regions of a linear demand curve.

equation to give

$$D = KP^{\varepsilon}$$

where K is independent of price. This demand function, due to its property of constant elasticity at every price point, is called the *constant elasticity demand function* or the *isoelastic demand function*.

Thus, using log–log transformed demand and price data, the estimate of the slope coefficient provides an empirical estimate of elasticity. The data could come from various sources – sales data, panel data, conjoint studies, surveys, and experimental methods have all been used. Econometrically, it may be necessary to control for such things as competitive and product-line prices, endogenous price, heterogeneity, and correlation with other marketing mix variables to improve the validity of the model. Often with multibrand, multisegment demand models, the elasticity is neither a model parameter nor amenable to derivation from the estimated parameters. Then, after estimating the demand function, elasticities for each brand and segment can be numerically evaluated at a few different price points and an average value reported.

In the preceding discussion, the dependent variable is demand or sales, but it could also be the market share. The price variable can also be variously defined, for example, as absolute or relative price, regular or promotional price, or unit price. There is some discussion that promotional price elasticities are higher than regular price elasticities (Blattberg, Briesch, and Fox, 1995). The former is calculated by adding a discount variable and/or a dummy variable to indicate whether the purchase was made when the product was on promotion. As another operationalization issue, elasticity, by definition, should be unchanged whether computed from an incremental increase or an incremental decrease in the price, but the two can give slightly different results due to reference price effects. The effects of different operationalizations and additional factors on the magnitude of elasticity have been studied. Tellis (1988) and Bijmolt, Van Heerde, and Pieters (2005) provide meta-analyses of a large number of econometric studies. In addition, an excellent discussion of empirical generalizations from elasticity and cross-elasticity research is provided in Hanssens, Parsons, and Schultz (2001, Ch. 8).

Optimal Pricing

Having understood how elasticity is defined and measured, we now turn to the implications of elasticity for pricing beginning with the monopoly setting.

A monopoly firm's profit from selling its product at price P when facing a demand function $D(P)$ and a cost function $C(D(P))$ is $PD(P) - C(D(P))$, that is, the total revenue minus the total cost. The firm's objective is to maximize its profit with respect to price. We assume a downward sloping demand curve and a convex, increasing cost function, both twice differentiable, and denote $D' = \partial D(P)/\partial P$ and $C' = \partial C(D)/\partial D$. The latter is called the *marginal cost* because it is the increase in total cost when one additional unit of the product is sold, or equivalently it is the cost of the last unit sold. The first-order condition for maximum is $(P - C')D' + D = 0$.

If the second-order condition is also satisfied then, from the definition of elasticity, substitute D' by $D\varepsilon/P$ into the first-order condition and it can be rewritten as

$$\frac{P - C'}{P} = \frac{1}{|\varepsilon|} \tag{3}$$

This is sometimes called the *inverse elasticity rule* for pricing. The left-hand side, which is the markup as a proportion of the price, is called the *Lerner Index*. It is a measure of the firm's market power, that is, the ability of the firm to set a price higher than its marginal cost. In comparison, under perfect competition, the firm has no market power and sets price at marginal cost. The right-hand side of the equation is the inverse of the elasticity at that price point. One can conclude that if a product has more elastic demand then, all else being equal, the optimal price will be lower.

Because the Lerner Index is bounded between 0 and 1, the right side of the equation must also be similarly bounded. It follows that $|\varepsilon| \geq 1$. Thus, when faced with inelastic demand, the firm should price higher till the demand becomes elastic. The optimal monopoly price lies in the

elastic portion of the demand curve. This result can also be argued as follows: total revenue is $PD(P)$ and the impact on it of a small increase in price is $PD' + D$, which may be rewritten as $D(1 - |\varepsilon|)$. This impact is 0 for unit elastic demand (implying that revenue is maximized at $|\varepsilon| = 1$), positive for inelastic demand, and negative for elastic demand. With inelastic demand, therefore, price can be increased to generate additional revenue. Thus, governments can more heavily tax goods such as gasoline and alcohol that have more inelastic demand to raise revenues.

Optimal pricing under Cournot competition. Next, we examine a model of imperfect competition. Suppose there are N competing firms indexed by $i = 1 \ldots N$. The total production in the industry is denoted by $Q = \sum_{i=1}^{N} Q_i$, where firm i produces quantity Q_i. Under Cournot competition, the market price $P(Q)$ is a function of the industry output. Each firm has the objective of maximizing its profit by its choice of output, that is, $\max_{Q_i} P(Q)Q_i - C_i(Q_i)$. The profit maximizing condition for firm i can be written as

$$\frac{P(Q) - C_i'(Q_i)}{P(Q)} = \frac{s_i}{|\varepsilon|} \tag{4}$$

where, $s_i = Q_i/Q$ is firm i' s market share. This equation is an extension of the inverse elasticity rule from the last section. It states that the larger the market share of a firm, and the lower the elasticity, the higher will be its markup.

Aggregating the left-hand side of the equation over all firms in the industry shows how far the industry deviates from perfect competition wherein the markups would be 0. Multiplying both sides of the equation by s_i and summing over firms yields

$$\sum_{i=1}^{N} \frac{P(Q) - C_i'(Q_i)}{P(Q)} s_i = -\frac{1}{\varepsilon} \sum_{i=1}^{N} s_i^2 = \frac{HHI}{|\varepsilon|} \tag{5}$$

where, $HHI = \sum_{i=1}^{N} s_i^2$ is the Hirschman–Herfindahl Index. This index is used by antitrust entities, for example, to assess the impact of a merger. In a market with a large number of symmetric competitors, the HHI will be close to 0.

The analysis in this section is useful for explaining regular price levels based on the elasticity. Firms would like their demand to be less elastic so that they can charge a higher price without causing a steep drop in demand and thereby improve their profitability. We therefore proceed to examine what factors influence elasticity and how marketers might apply them. Discussions on this topic can also be obtained from pricing textbooks such as Monroe (2003) and Nagle and Hogan (2006).

Determinants of Elasticity

The demand curve is an aggregation of individual customer's demands and elasticity is a measure of their price sensitivity. Customers will be more price sensitive, that is, more likely to reduce or stop their purchases of a product whose price has increased if they perceive that there are other, attractive substitutes available. Since the perception of substitutability is influenced by such things as customer satisfaction, differentiated positioning, and advertising, elasticity can be a useful measure of the health of the brand.

Type of product. The availability of substitutes may be constrained by the nature of the product category. If the product category is a necessity, such as salt, gasoline, or pharmaceuticals, then it will possibly have inelastic demand because customers have little choice but to make purchases. For similar reasons, addictive product categories such as cigarettes and alcohol tend to exhibit more inelastic demands. In the case of pharmaceuticals, some additional factors explain lower price sensitivity. For example insurance providers absorb part of the cost and furthermore pharmaceuticals, like textbooks, are prescribed by decision makers who do not bear the cost of the purchase.

Consumer characteristics. The availability of substitutes is also influenced by the characteristics of the customers, which may be influenced by targeting decisions. Elasticity can vary by customer segment depending on the substitutes available to each segment. For example, demand for restaurant food is less elastic for tourists than for locals because the former do not have

information about alternatives, the time or inclination to search, or the option to eat at home. A similar contrast exists between loyal shoppers and switchers, where mainly the latter look around for deals. Differentiation on search costs (e.g., retirees vs employed), inventory costs (e.g., home ownership vs not), and transportation costs (e.g., car owners vs others) also exist and they influence the availability of deals to different segments. The firm should examine the possibility of price discrimination between segments with different elasticities.

Influence of marketing mix. Advertising can be used to affect the price elasticity. If advertising serves to inform consumers about substitutes or attract new consumers who are more price sensitive, then its overall effect is to increase elasticity. But when advertising is used to build loyalty, thereby reducing the need to search for substitutes, it should result in lower elasticity. The loyalty building role of advertising is often contrasted with the immediate sales response of promotions. There is some evidence that offering promotions can make consumers more price sensitive over time because they learn to wait for deals. In effect, the firm provides the expectation of a future, cheaper substitute and hurts its sales at regular prices. However, in the case where promotions lead to habit formation, the result should be reduced elasticity.

Scope of market. The magnitude of elasticity will depend on how broadly or narrowly the product category is defined. An analysis of food and manufacturing industries by Pagoulatos and Sorensen (1986) found that industry price elasticities were all in the inelastic range, some examples being meat packing plants (−0.7), fluid milk (−0.17), soft drinks (−0.05), cigarettes (−0.75), roast coffee (−0.12), canned fruits and vegetables (−0.23), and pet food (−0.10). However, marketers operating at the brand level very often face elastic demand. For example, a given brand of table salt can have elastic demand due to consumers switching between brands despite the fact that the category demand remains inelastic.

Variation over time. By the "long run" we mean that consumers have fully adjusted to a price change. Elasticities will often be higher in the long run than in the short run because more substitutes become available. For example, when the price of gasoline increases, consumers have little immediate choice except to continue purchasing it in the same quantities. In the long run, however, they can respond to the higher price by purchasing more fuel efficient cars. Another issue of dynamics is how price elasticities vary over the product life cycle. This is complicated because not only does the product change over time but so do the competitive forces and the consumer mix at each stage of the product life cycle. If later consumers are more price sensitive, elasticity increases, but if they are more loyal then elasticity decreases. The latter point was argued by Simon (1979) and finds support in Bijmolt, Van Heerde, and Pieters's (2005) meta-analysis.

Share of budget. Products whose purchases have a smaller impact on the household budget have less elastic demand. Consistent with this, Bijmolt, Van Heerde, and Pieters (2005) conclude that the magnitude of elasticity is consistently higher for durable goods, which are typically large ticket items, than for groceries. The size of the household should also matter in that large households should be more price elastic than smaller households for products such as groceries because the purchases are a larger proportion of their income. This hypothesis is supported in Hoch *et al.* (1995).

Cross Elasticity of Demand

A related concept to the (own-price) elasticity of demand covered so far is the cross-price (or cross-) elasticity of demand. The cross-elasticity of demand is a measure of the sensitivity of the demand of a product to changes in the price of other products in the market. The cross elasticity of the demand of product A to a change in price of product B is defined as

$$\eta_{AB} = \frac{\partial D_A / D_A}{\partial P_B / P_B} = \frac{P_B}{D_A} \times \frac{\partial D_A}{\partial P_B} \approx \frac{\%\text{ change in } D_A}{\%\text{ change in } P_B} \tag{6}$$

where $D_A(P_A, P_B)$ is the demand for product A and P_B is the price of product B.

Note that η_{AB} is not identical to η_{BA}. For example, an analysis of grocery store products found that brands could take significant market share away from other brands in the same price-quality tier or in lower tiers through price reductions, but not from brands in the tiers above themselves (Blattberg and Wisniewski, 1989). Sethuraman, Srinivasan, and Kim (1999) found that the cross-elasticities were highest for brands within the same tier.

If customers view two products A and B as substitutable, then an increase in the price of B makes the purchase of product B less attractive and increases the demand for its substitute. Thus, positive cross-elasticity implies that the two products are substitutes and a higher value of cross-elasticity means that they compete strongly. Similarly, negative cross-elasticity means that the products are complements, while zero cross-elasticity means that the products might not be competing for the same customers. Which products compete with each other, and how strongly, is part of *market structure analysis*. Example 3 illustrates the usefulness of cross-elasticities for market structure analysis.

Example 3. For the three brands A, B, and C in a market, what elements of market structure are revealed from Table 2?

(i) Positive cross-elasticities means the brands are substitutes and the magnitudes suggest that A and B compete more strongly with each other than with C. (ii) The cross-elasticities are asymmetric, with brand A less sensitive to the other brands' prices than they are of its. To put this quantitatively, for brand i, *clout* is defined as $\sum_j \eta_{ji}^2$ and *vulnerability* as $\sum_j \eta_{ij}^2$ (Kamakura and Russell, 1989). Brands A, B, and C have clouts {0.4, 0.09, 0.0} and vulnerabilities {0.09, 0.36, 0.05} respectively. (iii) Cross elasticities are much smaller than elasticities. This means that additional sales are obtained either from stockpiling or market or volume expansion, not just from substitution. (iv) The market is not based on logit competition which generates uniform cross-elasticities. But, as in the logit, it is possible that brand A has the highest market share given that it has the lowest elasticity and highest cross-elasticities. □

The question of how much of the sales increase caused by a promotion can be attributed to purchase acceleration, quantity expansion, or brand switching is very relevant for retail store promotions because if the increase in sales comes mainly at the expense of competing brands in the store, then the promotion may not be worthwhile for the retailer overall. To explore this issue, Gupta (1988) proposed the following method for decomposition of elasticity. The sales in units of brand j on a shopping trip is stated as

$$\underbrace{D_j}_{\substack{\text{Sales of}\\ \text{brand } j}} = \underbrace{\Pr(I)}_{\substack{\text{Prob. of purchase}\\ \text{in category}}} \times \underbrace{\Pr(C_j|I)}_{\substack{\text{Prob. of choosing}\\ \text{brand } j \text{ given}\\ \text{category incidence}}} \times \underbrace{Q_j}_{\substack{\text{Quantity bought}\\ \text{given purchase}\\ \text{of brand } j}}$$

Differentiating with respect to price, the elasticity of sales of brand j is $\varepsilon_j = \varepsilon_{I_j} + \varepsilon_{C_j} + \varepsilon_{Q_j}$. For coffee data, Gupta (1988) found $\varepsilon_{I_j}/\varepsilon_j = 0.14$ (a measure of purchase acceleration), $\varepsilon_{Q_j}/\varepsilon_j = 0.02$ (a measure of quantity acceleration), and $\varepsilon_{C_j}/\varepsilon_j = 0.84$ (a measure of brand switching). Additional commentary on the interpretation of these numbers is provided by Van Heerde, Gupta, and Wittink (2003).

In addition to explaining competitive interactions, cross-elasticities are helpful for pricing a product-line sold by a firm. In the case of a two-product product-line with constant marginal costs c_i, the firm maximizes

$$\max_{P_1,P_2} \sum_i (P_i - c_i) D_i(P_1, P_2) \tag{7}$$

The results suggest that if the product-line is composed of complementary (respectively, substitute) products, the optimal product line prices will be lower (respectively, higher) than the optimal prices in isolation. Simon (1989) and Rao (1993) have formulae that incorporate cross-elasticities for multiproduct pricing.

Applying this type of pricing analysis in practice requires estimating the matrix of elasticities and cross-elasticities. Reibstein and Gatignon (1984), for example, estimate different cross-elasticity formulations to explain the pricing of eggs. Hoch *et al.* (1995) studied many more grocery product categories. Montgomery

Table 2 Matrix of elasticities and cross-elasticities.

	A	B	C
A	$\varepsilon_A = -2.2$	$\eta_{AB} = 0.3$	$\eta_{AC} = 0.05$
B	$\eta_{BA} = 0.6$	$\varepsilon_B = -2.8$	$\eta_{BC} = 0.01$
C	$\eta_{CA} = 0.2$	$\eta_{CB} = 0.1$	$\varepsilon_C = -3.3$

and Rossi (1999) suggest a Bayesian approach for estimating cross-elasticities.

Income Elasticity of Demand

In addition to prices, demand can be a function of other variables. In this section we consider the impact of income. The constant elasticity demand curves for a duopoly, for example, could be given by $D_i(P_i, P_j, Y) = K_i P_i^{\varepsilon_i} P_j^{\eta_{ij}} Y^{\kappa_i}\ \forall i, j \in \{1, 2\},\ i \neq j$, where Y is a measure of income and other terms are as previously defined. The income elasticity of demand D is defined by

$$\kappa = \frac{\partial D/D}{\partial Y/Y} = \frac{Y}{D}\frac{\partial D}{\partial Y} \approx \frac{\%\text{ change in demand}}{\%\text{ change in income}} \tag{8}$$

If income elasticity is positive, it means that the consumption of the product increases when income increases, and the good is called a *normal good*. For food items such as vegetables, and other necessities, it is observed that $0 < \kappa < 1$, that is, the proportion of budget spent on food decreases as income increases. This is called *Engel's Law*. In contrast, discretionary and luxury goods associated with leisure, recreation, and foreign travel have $\kappa > 1$. Finally, if demand decreases when income rises, the good is called an *inferior good*. For example, if, when income increases, consumers switch from eating potatoes to eating meat, then potatoes are an inferior good.

There is a connection between own elasticity, cross-elasticity, and income elasticity (Mas-Colell, Whinston, and Green, 1995, p. 27). Assume that if prices and income change by the same percentage then the consumer's demand for different goods does not change. Then for any i of the L goods in the market,

$$\varepsilon_i + \sum_{j \neq i} \eta_{ij} + \kappa_i = 0.$$

Acknowledgements

This write up benefited from input from Ali Bakhtiari, Octavian Carare, Sungha Jung, Vithala Rao, Raj Sethuraman, and César Zamudio, to all of whom I am very grateful. Errors are my own.

Bibliography

Bartels, R. (1976) *The History of Marketing Thought*, 2nd edn, Grid, Columbus.

Bijmolt, T.H.A., Van Heerde, H.J., and Pieters, R.G.M. (2005) New empirical generalizations on the determinants of price elasticity. *Journal of Marketing Research*, **42**, 141–156.

Blattberg, R.C., Briesch, R., and Fox, E.J. (1995) How promotions work. *Marketing Science*, **14**, G122–G132.

Blattberg, R.C. and Wisniewski, K.J. (1989) Price-induced patterns of competition. *Marketing Science*, **8**, 291–309.

Gupta, S. (1988) Impact of sales promotions on when, what and how much to buy. *Journal of Marketing Research*, **25**, 342–355.

Hanssens, D.M., Parsons, L.J., and Schultz, R.L. (2001) *Market Response Models: Econometric and Time Series Analysis*, 2nd edn, Kluwer, Boston.

Hoch, S.J., Kim, B.D., Montgomery, A.L., and Rossi, P.E. (1995) Determinants of store-level price elasticity. *Journal of Marketing Research*, **32**, 17–29.

Houthakker, H.S. and Taylor, L.D. (1970) *Consumer Demand in the United States: Analysis and Projections*, 2nd edn, Harvard University Press, Cambridge.

Kamakura, W.A. and Russell, G.J. (1989) A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research*, **26**, 379–390.

Mas-Colell, A., Whinston, M.D., and Green, J.R. (1995) *Microeconomic Theory*, Oxford University Press, New York.

Montgomery, A.L. and Rossi, P.E. (1999) Estimating price elasticities with theory based priors. *Journal of Marketing Research*, **36**, 413–423.

Monroe, K.B. (2003) *Pricing: Making profitable Decisions*, 3rd edn, McGraw-Hill/Irwin, Boston.

Nagle, T.T. and Hogan, J. (2006) *The Strategy and Tactics of Pricing: A Guide to Growing More Profitably*, 4th edn, Prentice-Hall, Upper Saddle River.

Pagoulatos, E. and Sorensen, R. (1986) What determines the elasticity of industry demand. *International journal of Industrial Organization*, **4**, 237–250.

Rao, V.R. (1993) Pricing models in marketing. Chapter 11, in *Handbooks in OR and MS*, vol. 5 (eds J. Eliashberg and G.L., Lilien), Elsevier, pp. 517–552.

Reibstein, D.J. and Gatignon, H. (1984) Optimal product line pricing: the influence of elasticities and cross-elasticities. *Journal of Marketing Research*, **21**, 259–267.

Sethuraman, R., Srinivasan, V., and Kim, D. (1999) Asymmetric and neighborhood cross-price effects: some empirical generalizations. *Marketing Science*, **18**, 23–41.

Simon, H. (1979) Dynamics of price elasticity and brand life cycles: an empirical study. *Journal of Marketing Research*, **16**, 439–452.

Simon, H. (1989) *Price Management*, North-Holland.

Tellis, G.J. (1988) The price elasticity of selective demand: a meta-analysis of econometric models of sales. *Journal of Marketing Research*, **24**, 331–341.

Van Heerde, H.J., Gupta, S., and Wittink, D.J. (2003) Is 75% of the sales promotion bump due to brand switching? No, only 33% is. *Journal of Marketing Research*, **40**, 481–491.

direct and interactive marketing

Frank J. Mulhern

Direct marketing consists of marketing communications where organizations communicate with consumers or businesses through direct media. Direct media include mail, print catalogs, and telephone. Interactive marketing is an extension of direct marketing into electronic media technologies that provide immediate, two-way communications. Interactive media include e-mail, the Internet, and mobile (wireless) marketing. Direct and interactive marketing can be contrasted with mass media, which communicate with large audiences, and personal selling, which communicates through person-to-person, nonmediated communications.

Compared to mass marketing, direct and interactive marketing allow organizations to communicate with consumers through conversation-like exchanges that generate integrated marketing communications (*see* INTEGRATED MARKETING COMMUNICATION STRATEGY) across media platforms (Shepard, 1999).

The key dimensions of direct and interactive marketing are as follows:

- *Information-driven*: Direct and interactive communications are typically linked to customer databases containing information on individual customer purchase histories. Customer databases facilitate three key aspects of direct and interactive marketing:
  - *Precise targeting*: The selection of target customers based on the analysis of past purchase behavior.
  - *Customization*: The ability to vary information, messages, prices, or promotions across customers.
  - *Addressability*: The use of the targeted customers' names within communications to achieve personalization.
- *Interactivity*: Interactive media allow for two-way communications between buyers and sellers as exemplified by e-mail and e-commerce web sites. Interactive communications feature the properties of conversations and facilitate ongoing interactions and relationship building.
- *Response mechanism*: Most direct marketing is direct response marketing – communications that contain a *call to action* – a specific request for a behavioral response such as making a purchase.
- *Measurability*: When paired with customer databases, the effects of direct and interactive marketing can be precisely measured, allowing for the evaluation of performance metrics including response rates, incremental sales, and return on investment.

One of the most important components of direct and interactive marketing is the use of information in databases to measure the buying behavior of customers and their value to the firm. As such, direct and interactive marketing blend with CUSTOMER RELATIONSHIP MANAGEMENT to evaluate key metrics of customer behavior including *recency* – how long ago a customer last purchased, *frequency* – the number of times a customer purchased within a time period, and *monetary value* – the dollar value of total purchases in that time period (Hughes, 2005). Database marketers (*see* DATABASE MINING AND MARKETING) use scoring models that combine such information into a single metric to prioritize who should be targeted. Direct and interactive marketing tactics are then used to solicit response from the consumers with the highest scores. The combination of direct and interactive media with customer databases allows firms to carry out a *long-term contact strategy* – the strategic management of communications with individual consumers over time based on the estimated customer lifetime value (Venkatesan and Kumar, 2004; *see* CUSTOMER LIFETIME VALUE (CLV)).

As media become more digital, the direct and interactive marketing approach becomes central to marketing (Harden and Heyman, 2009). In short, mass marketing approaches pioneered in the print and broadcast media era yielding to more precise and efficient marketing communications driven by information in customer databases and the capabilities of interactive media technologies. As television becomes digital and merges with the Internet, the direct and database marketing model will dominate advertising and bring about more

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

complete MULTICHANNEL MARKETING. In many ways, direct and database marketing represent a revolution in marketing practices that feature a high level of information and the ability to communicate interactively with individual customers.

The limitations of direct and database marketing stem from intrusiveness and customer privacy. Direct and interactive media can be very bothersome to consumers who receive an abundance of unwanted telephone calls, mails, e-mails, and online ads. Privacy concerns are the most serious challenge to direct and interactive marketing (Peltier, Milne, and Phelps, 2009). Many consumers oppose the collection and use of extensive information on their purchase behavior. Direct marketers, in conjunction with the Direct Marketing Association, facilitate mechanisms to protect consumer privacy. A common practice, particularly for mobile direct marketing, is to limit direct marketing to consumers who opt-in to solicitations from specific brands.

Bibliography

Harden, L. and Heyman, B. (2009) *Digital Engagement: Internet Marketing that Captured Customers and Builds Brands*, AMACOM, New York.

Hughes, A. (2005) *Strategic Database Marketing: The Masterplan for Starting and Managing a Profitable, Customer-Based Marketing Program*, McGraw-Hill, New York.

Peltier, J.W., Milne, G.R., and Phelps, J.E. (2009) Information privacy research: framework for integrating multiple publics, information channels and responses. *Journal of Interactive Marketing*, **23** (2), 191–205.

Shepard, A. (1999) *The New Direct Marketing: How to Implement a Profit-Driven Database Marketing Strategy*, McGraw-Hill, New York.

Venkatesan, R. and Kumar, V. (2004) A customer lifetime value framework for customer selection allocation strategy. *Journal of Marketing*, **68** (4), 106–121.

e-commerce and Internet marketing

Adel I. El-Ansary

The advent of the Internet transformed traditional, localized "market places" into vast contemporary global "market spaces." The transition has been enabled by features that characterize the Internet such as digitization, interconnectivity, interactivity, global reach, time-free and location-free access, personalization and customization, and market deconstruction. Market deconstruction is of particular significance because it leads to the creation of infomidiaries that replace and/or supplement the traditional "brick and mortar" marketing channels. These features singlehandedly or in combinations, facilitate the effective and efficient formulation and implementation of business strategy in general and marketing strategy in particular. Examples include (i) market segmentation and targeting through the creation of Internet "communities" powered by blogs and social media, (ii) one-on-one customer reach and interactive buying and selling through "e-commerce" portals, (iii) differentiation and positioning through "online advertising" and enabling customization and personalization on Internet portals, (iv) "online branding" (*see* BRAND COMMUNITY), and (v) marketing program implementation involving online catalog merchandizing, advertising, sales promotion, public relations, and the delivery of digitized products such as music, books, and software. In essence, the Internet established the foundation for building "new, narrowcast media" that enables firms to target customers through one-to-one communication.

E-commerce, enabling buying and selling transactions via the Internet, has shaped the way the world does business, making marketing more competitive, economic, and widespread. Businesses depend on the Internet to reach customers around the world and to market their products to every geographic location with access to the Internet. Businesses are realizing more than ever before that the customer's priority is the value (*see* CUSTOMER SATISFACTION) of a product. The Internet enables value creation through "portal content," "e-enabled" customer service, and "e-enabled" customer relations. The Internet also facilitates communication of value through an "e-integrated marketing communication mix" that includes (i) three-dimensional "online catalog merchandizing," (ii) online advertising that reaches targeted viewers around the world at a fraction of the cost of the "shotgun" advertisements in magazines, newspapers, and other media, (iii) online sales promotions designed to attract new business with discounts, coupons, and sweepstakes, and (iv) online publication of newsletters and other public relations tools.

The arrival of the Internet shifted power from the marketers to the consumers. Consumers have the ability to search the web for the best deals and products. Consumers are not bound by their geographic location or distance from the seller. The Internet provides unlimited information to consumers and grants them the ultimate purchasing power. Customers enjoy the ability to research information (*see* KNOWLEDGE ACCESSIBILITY) about the product on the Internet, compare prices, and make purchasing decisions. They can access information 24 hours a day, 7 days a week and are thus no longer limited by brick and mortar store hours. Prospective consumers can even view inventory and ship products to their doorstep. One of the greatest conveniences afforded by e-commerce and the Internet is the time savings for the consumer. Customers no longer have to search for a product and waste precious time comparing prices. With a single search on the Internet, consumers can view all available products and compare prices. Also, marketers can easily gather information from customers and process such information using countless analytical tools (*see* WEB SURVEYS) to better understand buyer attitudes, motives, consumption patterns and buying decision making, and behavior. Such analysis is essential for marketers to effectively and efficiently reach buyers. Consequently, marketers are more aware than ever before of what consumers' *desire*. (*see* CONSUMER DESIRE)

E-commerce and Internet marketing brings additional benefits and revenue to the firm. Businesses can communicate with customers, via email. They can answer customer questions, mitigate buyer concerns, and address problems

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

more effectively. Customers can provide feedback to businesses, allowing companies to gain a better grasp of their performance and *learn* (*see* CONSUMER EXPERTISE) how to improve the way they do business. In light of all these benefits, some firms only sell their products online (*see* ONLINE CONSUMPTION), and therefore have decreased sales costs and customer service costs. Because people around the world are able to view the products that a business offers (*see* DIGITAL MEDIUM AND GLOBAL MARKETING), a broader market is reached and more sales transactions are likely to occur.

E-business has not always been accepted; however, companies around the globe are realizing, more than ever before, that e-commerce is essential to their survival and their ability to thrive in a globally competitive market space. It is becoming a truism that the Internet is an equalizer of businesses, large and small, in countries, developed and underdeveloped. In essence, the Internet aids in establishing an "information-based democracy" (*see* CONSUMER WELL-BEING), enabling participation by everyone everywhere.

Bibliography

Chaffey, D. (2008) *E-Business and E-Commerce Management*, 3rd edn, Prentice-Hall.

Laudon, K. and Traver, C.G. (2009) *E-Commerce*, 5th edn, Prentice-Hall.

Plunkett, J. (2009) Plunkett's E-Commerce and Internet Business Almanac 2009. Plunkett Research.

Strauss, J., El-Ansary, A., and Frost, R. (2006) *E-Marketing*, 4th edn, Prentice-Hall.

ethical marketing and marketing strategy

Gene R. Laczniak and Patrick E. Murphy

Introduction

Ethical marketing can be defined as practices that emphasize "transparent, trustworthy, and responsible personal and/or organizational marketing policies and actions that exhibit integrity as well as fairness to consumers and other stakeholders" (Murphy *et al.*, 2005). This definition reflects the *normative* dimension of marketing ethics; that is, it embodies an aspirational view of how marketers *ought to* oversee their actions, programs, and procedures as they strategically engage their target markets. The approach is consistent with the motivational spirit of professional codes for marketing practitioners such as the American Marketing Association (AMA) Statement of Ethics (2008), which asks members to commit themselves to embracing "the highest standards of professional norms and values implied by our responsibilities to multiple stakeholders . . ."

Of course, like with most statements of ideals, the difficulty and debate occurs in the details and interpretation of complex concepts such as *transparency*, *integrity*, creating trust, and determining the meaning of *fairness* in particular exchange situations. In the points that follow, we try to make more pragmatic and tangible the conceptual dimensions that underscore the essentials of ethical marketing strategy. This is done in three ways. First, we distill the notion of normative ethical marketing into seven basic propositions (BPs) (Laczniak and Murphy, 2006) (referenced as LM, hereafter). Then, we link these propositions to actual company approaches that have been acclaimed as "best practices" in marketing ethics (Klein, Laczniak, and Murphy, 2006). Third, both of these perspectives will also be connected to the AMA Statement of Ethics (2008) in order to underscore the contention that there exist distinct norms and values to which professional marketers are expected to conform.

Some further words about this blended approach to clarifying the nature of ethical marketing are in order. In LM, the authors set about reviewing various frameworks in moral philosophy and models of corporate social responsibility (CSR) in order to derive the BPs that comprise ethical marketing. This is analogous to the promulgation of principles essential to informed marketing management such as "understand consumer needs," "segment markets according to meaningful criteria," and "measure customer satisfaction." Like most sets of marketing principles, the normative moral perspectives discussed in this article are "integrative" in that each informs and supplements the others.

Klein, Laczniak, and Murphy (2006) examine in detail the practices of over a dozen companies that have been designated as "ethical exemplars" either by winning the *Torch Award* of the U.S. Better Business Bureau or being so anointed in various case study analyses of role-model organizations. Thus, this second approach represents a "boots on the ground" description of how acclaimed organizations implement distinctive ethical marketing practices.

Finally, the AMA Statement of Ethics (2008) captures the fundamental norms and values that professional marketing practitioners should follow (See Table 1 – AMA Statement.) These values are explicated further into a "duty-based" set of expectations that professional marketers ought to adhere to. To the extent that the ethical precepts of the three approaches sketched above converge, we begin to build a strong case for the face validity of certain practices being essential to ethical marketing strategy. The following also must be stated: whether or not *being ethical* contributes to the profitability of marketing strategy is rather beside the point since the expectation is that marketers will practice their craft in conformance to legal requirements and ethical standards. That said, we believe that ethical marketing and improved long-term profitability are tethered as we have argued at length elsewhere (Murphy *et al.*, 2005). Indeed, as Aaker (2008) observes, "Strategy has to win not only in today's marketplace but in tomorrow's when the customer, the competitor set, and the market context may all be different." One way of generating such ongoing, strategic competitive advantage is for the firm to build an enduring reputation of trust in the marketplace that translates to brand loyalty and brand equity. With this in

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

mind, we unfold the elements that can help create such abiding trust in the marketing firm.

Ethical Marketing Puts People First

That "good" marketing is *person focused* is the first basic perspective of ethical marketing; it is referred to as the "*societal benefit*" BP (LM). Marketing theory has long placed effective marketing as something more than the *financially focused* view of maximizing profit for owner-investors. Moreover, marketing has always elevated the customer to a status coequal to that of shareholders. As marketing luminary Peter Drucker (1954) wisely opined, "There is only one valid definition of business purpose: to create a customer" (p. 37). Indeed the fundamental tenet underlying modern marketing has always been the marketing concept and its strategic implication that marketing planning is driven by the discovered needs and desires of consumers (Anderson, 1982). From that perspective it also follows that, strategically, organizational resources should be aligned in a manner that creates sustainable competitive advantage for the firm (Porter, 1998). As marketing strategy became more networked and outsourced (Dyer, Kale, and Singh, 2001), the recognition that other stakeholders were essential to marketing success became more prevalent. Today, the enlightened marketing organization realizes that (as Laczniak and Murphy, 2006 restate BP1), "persons in marketing transactions should never be viewed as merely a means to a profitable end" (p. 159).

This deep concern for the welfare of all parties to the marketing process can be seen in the practices of various acclaimed companies (Klein, Laczniak, and Murphy, 2006). For example, Charter Manufacturing Company of Mequon, Wisconsin uses a "cost-plus" pricing model that includes posting the indices from which price is derived on the Web for client-customers to better understand and monitor. Charter also gives its employees a great deal of autonomy both to do their own jobs and also to oversee their clients. And, Tom's of Maine (ToM), a company producing all-natural personal care products since 1970, includes the statement "People and nature have inherent worth and deserve respect" as part of their guiding "seven intentions" of "value-centered leadership". ToM had been named for 10 consecutive years as one of the "100 Best" companies to work for by *Working Mother* magazine.

The general sentiment of "ethical marketing puts people first" is also reflected in multiple passages of the AMA Statement of Ethics (2008) including the elaborations under the value of *respect* that marketers should strive to: "listen to the needs of customers and make every effort to monitor and improve their satisfaction . . ." and "treat everyone, including our competitors, as we would wish to be treated."

Ethical Responsibility Exceeds Legal Requirements

Ethical marketing seeks to achieve a behavioral standard in excess of the law. This is the second BP of ethical marketing as defined in LM. It should be understood that the legal system of developed economies is typically a "lagging" institution and, as a result, ethical perceptions of currently controversial marketing practices inevitability foreshadow the future of the law. As a result, marketers need to be proactive in identifying what these marketing-connected problems are and fix them quickly. Moreover, any corpus of professional knowledge, since it reflects its own common body of specialized information, ought to also include guidelines that articulate its unique ethical obligations. Embedded within this BP is the recognition of an implicit "social contract." This social contract, which takes the form of recognizing the considerable societal impacts of marketing practice (i.e., effective exchange, creation of markets, service innovation, and marketing jobs), also requires accepting substantial social responsibility that supersedes economic efficiency alone (see for instance, Dunfee, Smith, and Ross, 1999). In other words, as a major player in economic system, Marketing has the obligation to work to make that system better – to discharge its "duty of beneficence" as philosopher W.D. Ross (1930) might phrase it.

GSD&M, an advertising agency out of Austin, Texas founded in 1971, seems to be an example of a firm that willingly goes beyond the requirements of the law. Not only do they chisel their

Table 1 AMA Statement of Ethics (*AMA Statement of Ethics (revised 2008)*).

Ethical Norms and Values for Marketers
Preamble The American Marketing Association commits itself to promoting the highest standard of professional ethical norms and values for its members (practitioners, academics, and students). Norms are established standards of conduct that are expected and maintained by society and/or professional organizations. Values represent the collective conception of what communities find desirable, important and morally proper. Values also serve as the criteria for evaluating our own personal actions and the actions of others. As marketers, we recognize that we not only serve our organizations but also act as stewards of society in creating, facilitating and executing the transactions that are part of the greater economy. In this role, marketers are expected to embrace the highest professional ethical norms and the ethical values implied by our responsibility toward multiple stakeholders (e.g., customers, employees, investors, peers, channel members, regulators, and the host community).
Ethical Norms As marketers, we must: • Do no harm. This means consciously avoiding harmful actions or omissions by embodying high ethical standards and adhering to all applicable laws and regulations in the choices we make. • Foster trust in the marketing system. This means striving for good faith and fair dealing so as to contribute toward the efficacy of the exchange process as well as avoiding deception in product design, pricing, communication, and delivery of distribution. • Embrace ethical values. This means building relationships and enhancing consumer confidence in the integrity of marketing by affirming these core values: honesty, responsibility, fairness, respect, transparency, and citizenship.
Ethical Values Honesty – to be forthright in dealings with customers and stakeholders. To this end, we will: • Strive to be truthful in all situations and at all times. • Offer products of value that do what we claim in our communications. • Stand behind our products if they fail to deliver their claimed benefits. • Honor our explicit and implicit commitments and promises. **Responsibility** – to accept the consequences of our marketing decisions and strategies. To this end, we will: • Strive to serve the needs of customers. • Avoid using coercion with all stakeholders. • Acknowledge the social obligations to stakeholders that come with increased marketing and economic power. • Recognize our special commitments to vulnerable market segments such as children, seniors, the economically impoverished, market illiterates and others who may be substantially disadvantaged. • Consider environmental stewardship in our decision making. **Fairness** – to balance justly the needs of the buyer with the interests of the seller. To this end, we will: • Represent products in a clear way in selling, advertising and other forms of communication; this includes the avoidance of false, misleading, and deceptive promotion. • Reject manipulations and sales tactics that harm customer trust.

Table 1 (*Continued*).

Ethical Norms and Values for Marketers
• Refuse to engage in price fixing, predatory pricing, price gouging, or "bait-and-switch" tactics. • Avoid knowing participation in conflicts of interest. • Seek to protect the private information of customers, employees, and partners.
Respect – to acknowledge the basic human dignity of all stakeholders. To this end, we will:
• Value individual differences and avoid stereotyping customers or depicting demographic groups (e.g., gender, race, sexual orientation) in a negative or dehumanizing way. • Listen to the needs of customers and make all reasonable efforts to monitor and improve their satisfaction on an ongoing basis. • Make every effort to understand and respectfully treat buyers, suppliers, intermediaries and distributors from all cultures. • Acknowledge the contributions of others, such as consultants, employees, and coworkers, to marketing endeavors. • Treat everyone, including our competitors, as we would wish to be treated.
Transparency – to create a spirit of openness in marketing operations. To this end, we will:
• Strive to communicate clearly with all constituencies. • Accept constructive criticism from customers and other stakeholders. • Explain and take appropriate action regarding significant product or service risks, component substitutions, or other foreseeable eventualities that could affect customers or their perception of the purchase decision. • Disclose list prices and terms of financing as well as available price deals and adjustments.
Citizenship – to fulfill the economic, legal, philanthropic, and societal responsibilities that serve stakeholders. To this end, we will:
• Strive to protect the ecological environment in the execution of marketing campaigns. • Give back to the community through volunteerism and charitable donations. • Contribute to the overall betterment of marketing and its reputation. • Urge supply chain members to ensure that trade is fair for all participants, including producers in developing countries.
Implementation We expect AMA members to be courageous and proactive in leading and/or aiding their organizations in the fulfillment of the explicit and implicit promises made to those stakeholders. We recognize that every industry sector and marketing subdiscipline (e.g., marketing research, e-commerce, Internet selling, direct marketing, and advertising) has its own specific ethical issues that require policies and commentary. An array of such codes can be accessed through links on the AMA Web site. Consistent with the principle of subsidiarity (solving issues at the level where the expertise resides), we encourage all such groups to develop and/or refine their industry and discipline-specific codes of ethics to supplement these guiding ethical norms and values.

values into the foyer of their building (e.g., "Integrity – do the right thing") but they have turned down potential clients because partners at the firm were uncomfortable with what they were being asked to do. Similarly, Weber O'Brien, an accounting firm from Toledo, Ohio with 46 associates, has periodically terminated contracts when clients were found to be lying or cheating; they also offer substantially reduced fees to charitable organizations seeking accounting services (Klein, Laczniak, and Murphy, 2006).

Consistent with this BP, the AMA Statement of Ethics underscores the value of *citizenship*, which calls on marketers "to fulfill the economic,

legal, philanthropic, and social responsibilities that serve stakeholders" and includes the admonition to "contribute to the overall betterment of marketing and its reputation" (American Marketing Association, 2008).

Ethical Marketing Considers Intent, Means, and Ends

BP3 states that *marketers are responsible for whatever they intend as a means or end of a marketing action*. Further, if the intended means and end are acceptable, the action may proceed (with minor side-effects) unless there is risk of a major negative outcome for stakeholders. If such a major negative outcome occurs, even if it was not foreseen, marketers have a responsibility to try to make things right. LM explicate this idea by underscoring the obligation of all marketers to take responsibility for both how they conduct their marketing campaigns as well owning up to any negative social effects that result from their marketing activities, especially when these outcomes are foreseeable and substantial. For example, marketers of fast food and other high-fat/high-salt food snack sellers obviously are not singularly responsible for the obesity epidemic among children in the United States. Clearly, unbalanced eating habits, the lack of exercise, and too much time spent in front of computer or TV are major contributory factors; nevertheless, high calorie, possibly addictive junk food, is somewhere in the causal mix. Thus, according to this BP, marketers should be prudential about the manner in which they promote junk food and fast food and ought to take proactive steps to offer healthy alternatives as well as "smart-eating" advice.

Klein, Laczniak, and Murphy (2006) have profiled several firms that seem to go far beyond the black letter law in taking into account their responsibilities. For example, Toledo Metal Spinning, an industrial fabrications producer in business since 1929, totally lost their major production and operations facility in a 1998 fire. In an action reminiscent of the well-known Malden Mills case, they kept their employees fully engaged in salvage, cleanup, and reconstruction for seven months until a new facility was ready, all with an eye to making progress toward fulfilling the 500 orders that they had "in process" with an array of customers. Similarly, Baxter International (est. 1931), a global supplier of medical technologies and health care products, was quick to assume the product liabilities to injured patients from defective dialysis filters even though Baxter had only recently acquired the rights to the problematic product line from another company.

The AMA Statement of Ethics (2008) speaks of this disposition toward responsibility quite explicitly. The first *general* norm of the statement is "Do no harm," the most common ethical dictum across a variety of codes in the professions including medicine and engineering. And, under the value of *transparency* the AMA statement also reads as follows:

Explain and take appropriate action regarding significant product and service risks, component substitutions or other foreseeable eventualities that could affect customers and their perceptions of the purchase decision.

Ethical Marketers Try to Inspire Moral Imagination

The fourth BP focuses on categorizing marketing managers according to their level of moral reasoning. In most firms, managers making marketing decisions differ in their ability to evaluate and resolve ethical issues. Some marketing executives will have little ethical sensitivity while others will have the capacity for significant *moral imagination* – that is, the character and ability to morally reason to creative ethical solutions when encountering an ethical issue (Werhane, 1999). Life experiences, personal values, and basic character traits vary among marketing managers and influence their critical ethical evaluations. *Thus, organizations should seek to understand the nature of these different moral aptitudes and work to instill an improved ethical reasoning capacity among all their managers.*

Inspired by the work of Kohlberg (1969) and others, LM classify marketing managers into four major types detailed below.

- *Egoist marketing managers* are the least morally developed and have a strong tendency to resolve ethical situations on the basis of their own immediate interests and

consequences. Individuals at this comparatively undeveloped stage of moral thinking give strong weighting to the incentive and sanctions that will affect only them. They use rationalizations such as "everyone else is doing it" and "we were only following orders."

- *Legalist marketing managers* overtly espouse the law as their guide in adjudicating the propriety of any marketing action. As hinted at in BP2, they embrace predominately an agency approach to their managerial duties and generally follow the adage that "If it is legal, it is ethical."
- *Moral strivers* are those marketing managers who have progressed in their moral thinking and development to the point where they are capable of considering and balancing multiple stakeholder claims when deciding what constitutes an ethical imperative. Empathy for others is what distinguishes moral strivers from the lower levels. However, without sufficient reinforcement by company ethical standards and training, some strivers will fall back to legalist or egoist tendencies.
- *Principled marketing managers* have reached a high level of moral development. Managers who attain this sophisticated state address their ethical problems by regularly applying both prevailing ethical norms and applicable laws to the specific situation. They possess substantial moral imagination and bring this to bear on their decisions. Like Maslow's "self actualization" stage, few managers reach this highest level of development (p. 163).

While these four categories are rough approximations, one can see their relevance to marketing management. Recent examples of egoist marketing managers are those from Merrill Lynch and AIG who were "secretly" rewarded with bonuses. Almost 700 Merrill managers received bonuses of around $1 million while 14 individuals received a combined $249 million (Bray, 2009). Then in March of 2009, while the Federal government presumptively owned busted insurance company AIG, their traders (who had lost tens of billions of dollars with their earlier derivative-based trades) were given over $100 million in bonuses by management in order to incentivize their retention. Recognizing the frail state of the US economy and the fact that both Merrill and AIG were part of the US government bailout in 2008–2009, these episodes may supplant the vilified Enron managers as the personification of ethical egoism even as the term "corporate swine" embodied the lexicon of public outrage.

Somewhat on safer ground, *legalist* managers are quick to fall back on the law as the guiding force for all decisions. Generally, marketing managers in the building and construction sector are known for their letter-of-the-law tendencies perhaps because of their engineering background, a get-the-job-done mentality, and the process-dominated nature of that industry.

Most marketing managers fall into the *moral striver* category. Some of the recent strategic adjustments in the food industry, responding to critics who assert that consumer packaged goods companies are not sensitive enough to the growing obesity epidemic in the United States, are indicative of moral striving. One recent positive response comes from SuperValu, a major Midwestern supermarket chain, which is putting color-coded labels on its store brands to help consumers make a more informed choice regarding fat, sugar, calorie count, and other nutritional content.

While it is perilous to label all managers within a particular firm as "*principled*," two companies are used to illustrate this stage. The first is Honest Tea. The company was founded as a competitive response to the then typical flavored, cold teas that were generally high in sugar and low in nutrition content. The original (Honest Tea) product was made from whole tea leaves, which leads to better taste. The name fits this all-natural product because it strived to create healthy *and honest* relationships with multiple stakeholders – customers, suppliers, and the environment. (In fact, the original name was Honestea, but the threat of a lawsuit by Nestea forced the change). Furthermore, in order to enhance distributive justice in the supply chain, the company buys "fair trade" tea from around the world and partners with nonprofit organizations in selling some of its flavors. Honest Tea also invented and uses tea bags with a one-piece, tag-and-bag design and no staple or

string, making them fully biodegradable. The bag itself is made with unbleached fiber from an abacca plant (Klein, Laczniak, and Murphy, 2006). These principled policies led to financial success for Honest Tea, and in 2008, Coca Cola purchased a 20% stake in the firm.

The second company example is Lego, the well-known Danish toy brick producer. The name of the company means to "play well." Its managers have long followed principles such as only depicting toys on the box that can be assembled from the bricks inside and never having larger pictures on the outside of the box than "actual size" so that the children never will be misled. In recent years, Lego also has become a signatory and proponent of the UN Global Compact. The principles espoused there deal with child labor, health, and safety, and protection of the environment. Lego has also developed its own set of principles to guide its relationship with suppliers. While no company is perfect, Lego does seem to demonstrate in multiple ways that they are a "principle-based," ethically driven company.

The four types of marketing managers described above relate to the AMA Statement of Ethics in that the Association hopes that by following the ethical values outlined within, professional marketing managers will be pushed toward the "principled stage" of moral development. In fact, the "Preamble" to the American Marketing Association (2008) statement indicates that "marketers are expected to embrace the highest professional ethical norms and the ethical values implied by our responsibility toward multiple stakeholders . . ."

Ethical Marketers Embrace Core Values

With the formulation of BP5, LM suggest that "Marketers who aspire to operate on a high ethical plane should articulate and embrace a core set of ethical principles." There, five principles are recommended including the *principle of nonmalfeasance*, basically not intentionally doing harm via marketing actions, policies, and procedures, and the *principle of nondeception* – avoidance of intentional misrepresentation to or unfair manipulation of consumers. Both these precepts are reflective of the general norms (for good and fair marketing) specified in the American Marketing Association (2008) Statement of Ethics. In addition, three other principles are put forward in LM that challenge marketing practitioners to aspire to even greater levels of social responsibility. These are as follows:

- *The principle of protecting vulnerable markets*: Marketers must always take extraordinary care when engaging in exchanges with vulnerable [market] segments.
- *The principle of distributive justice*: There is an obligation on the part of all marketing organizations to assess the fairness of marketplace consequences flowing from . . . marketing practice.
- *The principle of stewardship*: Marketers are obligated to ensure that their operations will not impose external costs on society, especially the physical environment, that result from their . . . marketing [actions].

As documented by Klein, Laczniak, and Murphy (2006), there are multiple instances of organizations that attempt to formulate, embrace, and follow such core principles as part of their corporate mission and culture. For example, BadgerMeter, a Milwaukee-based manufacturer of water meters and flow technology that sells its products worldwide, has long held three guiding principles: (i) exemplary ethical conduct, (ii) respect for all people, and (iii) managing for the long term. These ideas are expanded in a detailed corporate code of conduct but interestingly, the organization specifies that all of its core vision can be realized by caring for each customer as if it were one of its "hometown" Milwaukee neighbors. Similarly, ToM, the producer and seller of natural and environmentally compatible personal-care and household products mentioned earlier, has articulated a "managing upside down" approach that begins with the assumption that "people and nature have inherent worth and deserve respect." This idea is at the heart of its "value-centered leadership" approach that encourages ToM managers to "connect with goodness" as the true path to economic success.

A particularly interesting example is bp, the energy sector multinational (i.e., the former British Petroleum), because it both illustrates

the importance of having core values as well as the reality that espousing these is not enough to insure ethical behavior. bp is an organization that, in the mid-2000s, experienced a horrific explosion at its Texas refinery resulting from employee inattention as well as pipeline damage in Alaska, due to its lack of following required safety procedures. That said, there was no doubt within the bp organization that what happened was an ethical failure because bp has taken great pains to document not only its core values (e.g., trust, acceptance of the UN Declaration of Human Rights) but also a decision-making framework that should be consulted along the way. This decision model, internally sometimes referred to as the "bp Way," makes clear that "there is no right way to do a wrong thing" and that any uncertainty about the ethicality of a contemplated decision requires greater consultation within the company.

Ethical Marketers Accept the Stakeholder Concept

The notion of a stakeholder (i.e., any group or individual who can affect, or is affected by, the achievement of the organization's objectives) was introduced to the business literature by Freeman (1984). Since that time, marketing managers and other senior executives have focused attention on those who have a "stake" in business decisions. It can be logically reasoned that *the adoption of a stakeholder orientation is essential to the advancement and maintenance of ethical decision making in all marketing operations*. Such an orientation embodies the idea that marketing organizations operate in and on behalf of society (Laczniak and Murphy, 2006). In the sixth BP, LM recognize three primary and three secondary stakeholders. Not surprisingly, customers, employees, and investors are the three primary stakeholder groups identified. These groups are primary because they are typically necessary to the completion of successful marketplace transactions and their claims trump other stakeholders. The three secondary stakeholders identified are distributors/suppliers (some of whom have contractual relationships with the marketing organization and are essential partners to the well-being of the firm) as well as host communities and the general public.

While many companies have not only embraced the stakeholder concept but also worked to genuinely "engage" in stakeholder dialogue, noted here are firms that have done a particularly superior job of addressing the needs of one of the six stakeholder groups. For example, customer orientation is a hallmark of many successful commercial marketers such as Johnson & Johnson. J&J's stakeholders include doctors, nurses, patients, mothers, and fathers and are given primacy in their famous Corporate Credo. Similarly, Harley Davidson – that famous international icon of customer admiration – is known worldwide for the brand loyalty and customer appreciation connected to that organization. In the nonprofit world, the Mayo Clinic is recognized as being unparalled in its patient orientation (Berry and Seltman, 2008). Perhaps, the best national example of focusing on the employee stakeholder is Southwest Airlines, a firm espousing the philosophy that "satisfied employees lead to satisfied customers." Also, Charter Manufacturing, identified earlier in this article, is a regional company illustration of a firm that accords its employees exceptional autonomy and respect. In contrast, a purely investor orientation is extremely well practiced by investment banks like Goldman Sachs – putting its investors (and some would say its management) above the needs of the market/consumers and sometimes taking advantage of its competitors in the process. The current state of the world's financial markets shows the catastrophic effect of only focusing on short-term investor stakeholders.

In the secondary stakeholder category, Toyota is known for its strong relationships with its suppliers and dealers. Similarly, Target and General Mills, headquartered in Minnesota, are renowned for responding with empathy to their (secondary) community stakeholders via their generous philanthropic policies. They and many other companies in that geographic area donate 5% of pretax profits to charity, much of it going to nonprofit corporations. The general public, as yet another secondary stakeholder, is often embodied by the media or advocacy groups. Two companies that have constructively engaged their secondary stakeholder critics during the last several years are Nike and Wal-Mart. In the late 1990s, Nike was widely vilified for

the working conditions of its subcontractors, especially those in the developing world. The company embarked on a major effort to negotiate with representatives of several human rights activist groups in order to improve the situation. A more recent illustration is Wal-Mart and its environmental critics. The firm has not only engaged these stakeholders, but made bona fide efforts to become a "green" company and to stock and sell products that are increasingly environmentally friendly (Rosenbloom and Barbaro, 1938). (This laudatory illustration pertains only to Wal-Mart's environmental efforts because its responses regarding employee compensation and supplier relations have been more controversial.)

In the AMA Statement of Ethics, "stakeholders" are cited numerous times. First, various stakeholder groups are specifically mentioned in the preamble. Second, stakeholders are explicitly referenced regarding every endorsed ethical value except *fairness* and one could argue that they are implicit in the notion of fairness, since that concept embodies giving relevant parties what is their due. Third, in the Implementation section, all AMA members are encouraged to be courageous and proactive in dealing with relevant stakeholders.

Ethical Marketers should Embrace a Process of Moral Reasoning

The final BP lays out a series of steps that marketing managers should follow in making an ethical decision. The first stage is cultivating ethical awareness and sensitivity. This is the domain of corporate culture and is reinforced by management. Marketers must be able to "see" an ethical issue when it arises. Without a sense of awareness and sensitivity, marketers are less likely to be able to ascertain whether a certain decision contains ethical implications. In one of the most famous articles in the business ethics literature, *The Parable of the Sadhu* (1983), Buzz McCoy writes about his mountain climbing experience in the Himalayas and pens a classic line that illustrates this point: "Real moral dilemmas are ambiguous, and many of us hike right through them, unaware that they exist."

The second step of the ethics decision-making protocol involves framing the ethical issue or question. It should be understood that the formulation of an ethical question does not imply that the questionable practice will necessarily be deemed unethical (Laczniak and Murphy, 2006, p. 169). It is the role of the marketing manager to assist junior members of the department in ascertaining whether a potentially questionable policy such as "product placement" or "ambush marketing" constitutes an ethical issue that needs discussion and resolution.

Articulation of stakeholders affected by the marketing decision in question is the third stage of ethical decision-making process. Both primary and secondary stakeholders should be specified. For example, in a decision about whether to introduce a new and/or more violent update to a video game, prospective customers (both adult and underage), employees of the firm, and company stockholders should be considered. Game developers and retailers who might distribute the software are also important secondary stakeholders. Because of the software's violent themes, both activist groups and society, in general, should be considered as well. If the company expects a potential public backlash, the secondary stakeholders and the fallout from the product launch needs to be addressed before moving forward.

The fourth step in the ethical reasoning process involves the selection of an ethical standard or standards. This could include simple maxims like the golden rule, the Wall Street Journal test or more complete ethical theories like utilitarianism, duty-based or virtue-based ethics. For example, the hypothetical decision of a local Ford dealer about whether to acquire a JB Byrider franchise should be evaluated using one or more of these standards. The Byrider franchise specializes in used automobiles that appeal to low to moderate income consumers. The past record of these franchises has been questioned because of the relatively high interest rates they charge and propensity to quickly repossess autos when customers fall behind on their payments (Grow and Epstein, 2007).

The fifth stage is "ethical analysis," which involves applying the ethical standard to the ethical question keeping in mind the impact on impacted stakeholders. Continuing with the Byrider example, the auto dealer may believe that acquiring such a franchise will lead to more benefits for consumers than costs. Hence, using

a utilitarian analysis, the decision to go forward might be positive. However, with most other ethical standards, the decision might be called into question because this branded franchise has not historically operated with transparency toward consumers (a violation of virtue ethics); moreover, it appears that Byrider intends to sell cars to at least some consumers who cannot afford them (i.e., possibly failing to meet the duty-based ethical standard). The point of this example is that if a manager is a "principled" one, s/he will use multiple ethical criteria to evaluate decisions.

Sixth, a decision should be made using all the information gathered in the preceding stages. Usually, there are three options: acceptable to go forward; the strategy may be ethical if amended in some fashion; or practice should be abandoned. A positive example of where a company moved forward with a vetted decision, after a thorough ethical evaluation process, is BzzAgent. Here the organization, a buzz marketing agency, moved from instructing its agents to be "discreet" to the higher standard of disclosure – "you must tell people" – when asked if they were involved in instigating a "word-of-mouth" marketing campaign.

The final stage of the decision protocol is to monitor the outcomes of the "ethical" decision. Steps need to be taken so that unintended consequences do not occur such as inadvertently using ultrathin models in advertising, thereby contributing to perceptions of poor body image among young girls. The highly acclaimed Dove campaign showing women of "many sizes and shapes" was a positive response to the long-standing trend of only using slender models in cosmetics advertising. Similarly, marketers who sell to children and other vulnerable consumers have a special obligation to monitor the societal effects of their advertising and product usage. Studying the impact of marketing actions on stakeholders beyond consumers is the type of monitoring that one might expect of managers who exhibit moral imagination.

Although the AMA Statement of Ethics does not articulate its own set of steps for making ethical decisions, the entire document is aimed at helping marketers to become more aware of the ethical issues. In fact, the three central ethical norms – do no harm, foster trust in the marketing system, and embrace ethical values – are helpful ethical guidelines that marketers can use in stage three and apply in stage four of the ethics assessment process. Finally, the six ethical values that embody the core of the AMA statement can be seen as an illustration of a "virtue ethics" approach to ethical issues.

Conclusion

Several conclusions can be drawn for marketing managers from the analysis of the seven BPs above. First, the marketing concept, strategy formulation, relationship marketing, and supply chain management are all inherently ethical activities. If any or all of these marketing practices are to be successful, a sense of trust and fairness must exist among marketers, their customers, suppliers, and end users of their products. This notion is implicit in the BPs. Clearly, the ethical dimensions of marketing are not new ideas but rather theory-inspired approaches that require renewed diligence.

Second, ethical issues facing marketers in the future will likely require both a normative and technical analysis. Most ethical problems are complex. One reaction by companies has been to hire ethics consultants or internal ethics officers to help them deal with these questions and set up ethics training programs. However, this is probably not enough any longer. For example, many environmental and solid-waste problems necessitate sophisticated technical advice on (for example) recycling versus landfill disposal. The "right" decision may not be clear until the scientific evidence is first factored in. Thus, both types of analyses should help illuminate the difficult choices ahead for marketing executives.

Third, marketing's role in society and in dealing with social issues should be examined from an ethical standpoint. Marketing as a business practice has made major contributions to society over time and occasionally is judged harshly from an ethical performance viewpoint. But marketing can be an exceptional force for "good." Marketing methods will increasingly be called upon to shape various social issues facing the world such as AIDS education, discouraging illegal drug use, helping publicize innovations in education, as well as various other social interventions that require promotion and

dissemination. Marketing has much to offer to the discussion and the potential alleviation of social problems; marketing-rooted strategy can contribute positively to environmental betterment and to meeting the needs of disadvantaged consumers in the future.

Fourth, ethical marketing companies exhibit a genuine stakeholder orientation and engagement. Exemplar companies, such as those profiled above, commit to represent all stakeholders because they recognize the rights, claims, and presumption that other affected parties expect of a business organization – especially recognitions that go beyond the basic requirements of the law. Our discussion of BP2 and BP6 directly illustrate this point. Although many firms list stakeholders as important in their website profiles, the most enlightened firms go the extra step in creating an ongoing dialogue with major stakeholders when changing conditions warrant such discussions.

Finally, ethical leadership is crucial for the success of any firm. While it starts with the CEO and CMO, other top-level marketing managers should also exhibit integrity and fairness in their dealings with employees and customers. While BP4 and BP7 both address the importance of managerial leadership, this idea pervades the very core of ethical marketing. The Implementation section of AMA Statement of Ethics calls on AMA members to be such leaders: "We expect AMA members to be courageous and proactive in leading and/or aiding their organizations in the fulfillment of the explicit and implicit promises made to those stakeholders." Although we have concentrated on companies rather than individuals in providing profiles of enlightened and ethical marketing strategy, we want to conclude by noting the reflections of two ex-CEOs who have written persuasively on ethics and values-based leadership – Bill George and Sims (2007) and Harry Kraemer (2007).

In the end, the purpose of marketing strategy is to develop assets and competencies that create "value propositions" appreciated by the firm's target markets on a sustained basis. The ethical perspectives discussed in the paragraphs above are both an articulation of how ethical marketing is conducted according to the highest professional ideals as well as a road map for the cultivation of expanded consumer trust that accrues to organizations that embrace these approaches.

Bibliography

Aaker, D.A. (2008) *Strategic Marketing Management*, 8th edn, John Wiley & Sons, Inc., New York.

American Marketing Association (2008) Statement of Ethics, accessed at http://www.marketingpower.com/ AboutAMA/ Pages/ Statement %20 of %20 Ethics.aspx (accessed on 2009).

Anderson, P.F. (1982) Marketing, Strategic planning and the theory of the firm. *Journal of Marketing*, **46**, 15–26.

Berry, L.L. and Seltman, K.D. (2008) *Management Lessons from the Mayo Clinic*, McGraw Hill, New York.

Bray, C. (2009) Merrill gave $1 Million Each to 700 of its Staff. The Wall Street Journal (February 12), Vol. 254, p. C3.

Drucker, P.F. (1954) *The Practice of Management*, Harper & Row, New York.

Dunfee, T.W., Smith, N.C., and Ross, W.T. Jr. (1999) Social contracts and marketing ethics. *Journal of Marketing*, **63** (2), 7–24.

Dyer, J.H., Kale, P., and Singh, H. (2001) How to make strategic alliances work. *Sloan Management Review*, **42** (4), 37–43.

Freeman, R.E (1984) *Strategic Management: A Stakeholder Approach*, Pitman, Boston.

George, B. and Sims, P. (2007) *True North: Discover Your Authentic Leadership*, Jossey-Bass, San Francisco.

Grow, B. and Epstein, K. (2007) The Poverty Business Inside U.S. Companies' Audacious Drive to Extract More Profits from the Nation's Working Poor. BusinessWeek (May 21), pp. 56–67.

Klein, T.A., Laczniak, G.R., and Murphy, P.E. (2006) Ethical marketing: a look on the bright side. *Marketing Management Journal*, **16** (1), 228–243.

Kohlberg, L. (1969) Stage and sequence: the cognitive developmental approach to socialization, in *Handbook of Socialization Theory and Research* (ed. D.A. Goslin), Rand McNally, Chicago, pp. 347–480.

Kraemer, H.M. Jr. (2007) What is this Thing called CEO Leadership? Directors and Boards, Fourth Quarter, 24–28.

Laczniak, G.R. and Murphy, P.E. (2006) Normative perspectives for ethical and socially responsible marketing. *Journal of Macromarketing*, **26** (2), 154–177.

Mc Coy, B.H. (1997) The parable of the sadhu. *Harvard Business Review*, **75**, 54–64.

Murphy, P.E., Laczniak, G.R., Bowie, N., and Klein, T.A. (2005) *Ethical Marketing*, Pearson Education, Upper Saddle River.

Porter, M. (1998) *Competitive Advantage: Creating and Sustaining Superior Performance*, The Free Press, New York.

Rosenbloom, S. and Barbaro, M. (2009) Green-light specials, now at Wal-Mart. The New York Times (January 25), p. 4–5. Ross, W.D. (1938) *The Right and The Good*, Clarendon Press, Oxford.

Werhane, P.H. (1999) *Moral Imagination and Management Decision Making*, Oxford University Press, New York.

first-mover (pioneer) advantage

Peter N. Golder

First-mover (or pioneer) advantage is the long-term reward that may accrue to the first company to enter a new market (*see* PRODUCT CATEGORY). Five key theories underlie a potential first-mover advantage.

- *Ease of recall*: People recall more easily the first brand they use or hear about in a new product category. As a result, pioneering brands can remain top-of-mind, promoting word-of-mouth and ongoing purchases (*see* CONSUMER MEMORY PROCESSES).
- *Brand loyalty*: Consumers often develop a preference for purchasing brands that they have used previously (*see* PERCEPTION OF BRAND EQUITY; CONSUMER BRAND LOYALTY). By continuing to purchase the same brand, consumers avoid the cost of searching for alternative brands as well as the risk of possibly sampling a poor brand. In new product categories, brand loyalty can have especially strong effects because first movers shape the expectations of the entire category. Essentially, first movers are prototypes against which later entrants are judged. First movers determine which attributes are important and which levels of attributes are appropriate. Thus, later entrants may be perceived as inferior to the standard set by the pioneer or be perceived as "me-too" copies of the original brand.
- *Technological leadership*: Pioneers can operate at the technological frontier of the new-product category. As a result, they have the opportunity to secure patents that may deter other firms from entering a category or at least inhibit their ability to compete (*see* INNOVATION TYPOLOGIES; INTELLECTUAL PROPERTY RIGHTS).
- *Economies of scale and experience*: First movers can be the first firm to achieve substantial volume in manufacturing and sales. This volume can result in lower costs and the associated experience may lead to higher quality and better features than the competitors' products.
- *Resource capture*: First movers have the opportunity to secure the best raw materials, suppliers, consumer segments, and distribution outlets (*see* MARKETING CHANNEL STRATEGY). Later entrants may be left with less profitable alternatives. For example, first movers can sell to innovators, who are eager to purchase new products, while later entrants are left trying to sell to noninnovators. In addition, first movers may develop broader product lines enabling them to appeal to a variety of market segments.

Many empirical studies report results that support a first-mover advantage on several key performance measures. First, they report an average MARKET SHARE for first movers of approximately 30%. Second, one contributing factor to this high market share is that remarkably few pioneers fail. However, early empirical studies do not report a specific failure rate. Third, first movers' high market shares result in the finding that most market leaders are first movers. Finally, studies report that market leaders tend to remain market leaders for many years, often decades.

Subsequent research has found that two key limitations result in first movers having substantially lower rewards than described above (*see* LATER MOVER (NONPIONEER) ADVANTAGE for more details). The first limitation is that initial studies include only surviving brands in a category. Therefore, the longest surviving brand may have been successful for a variety of reasons other than being the first mover. The second limitation in many studies is that companies classified themselves as either first movers or later entrants (*see* MARKET DEFINITION; SURVEY RESEARCH). As a result, successful firms were more likely to attribute their success to being a first mover on some aspect of their strategy even though earlier entrants employing inferior strategies may have failed.

The diminished rewards for first movers identified in later studies have led to a reconsideration of first movers' theoretical advantages. All of these theories rely on the presumption that the first mover was successful on one or more important dimensions. That is, the first mover offered a superior product, positioned it properly to the most attractive

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

market segment, locked up the most attractive distribution channel, and continued to benefit by staying at the forefronts of technology and experience. Some first movers are able to do these things, but most are not. As a result, companies seeking to establish a first-mover advantage would be wise to regard the situation as a first-mover opportunity instead. Long-term success will result from assiduously cultivating these opportunities over time.

Bibliography

Kalyanaram, G., Robinson, W.T., and Urban, G.L. (1995) Order of market entry: established empirical generalizations, emerging empirical generalizations, and future research. *Marketing Science*, **14** (3 Part 2 of 2), G212–G221.

Kerin, R.A., Varadarajan, P.R., and Peterson, R.A. (1992) First-mover advantage: a synthesis, conceptual framework, and research propositions. *Journal of Marketing*, **56**, 33–52.

Lieberman, M.B. and Montgomery, D.B. (1998) First-mover (dis)advantages: retrospective and link with the resource-based view. *Strategic Management Journal*, **19** (12), 1111–1125.

global marketing strategy

Johny K. Johansson

Introduction

A global marketing strategy (GMS) is a strategy that encompasses countries from several different regions in the world and aims at coordinating a company's marketing efforts in markets in these countries.

A GMS does not necessarily cover all countries but it should apply across several regions. A typical regional breakdown is as follows: Africa, Asia, and the Pacific (including Australia) Europe and the Middle East, Latin America, and North America. A "regional" marketing strategy is one that coordinates the marketing effort in one region.

A GMS should not be confused with a global *production* strategy. Outsourcing and foreign manufacturing subsidiaries, common features of a global production strategy, can be used with or without a GMS for the finished products.

As listed in Table 1, GMSs can involve one or more of several activities.

The coordination involved in implementing a GMS unavoidably leads to a certain level of uniformity of branding, of packaging, of promotional appeal, and so on (Zou and Cavusgil, 2002). This also means that a GMS, in some ways, goes counter to a true customer orientation (*see* MARKETING PLANNING). The product and marketing mix are not adapted to local preferences, as a customer orientation suggests. This is a potential weakness of GMSs, and leaves opportunities open for local products and brands.

Table 1 Components of a global marketing strategy.

Items Listed in Order of Descending Occurrence
• Identical brand names
• Uniform packaging
• Standardized products
• Similar advertising messages
• Coordinated pricing
• Synchronized product introductions
• Coordinated sales campaigns

As the notion of integrated marketing communications (*see* INTEGRATED MARKETING COMMUNICATION STRATEGY) suggests, the ensuing consistency can have positive revenue benefits because of reinforcement of a unique message, spillovers between countries, and so on. But the main driving force behind the adoption of a GMS is the scale and scope of cost advantages from such uniform marketing strategies. These cost advantages include elimination of unnecessary duplication of effort, savings on multilingual and same-size packaging, use of the same promotional material, quantity discounts when buying media, and so on. The pros and cons of a GMS are given in Table 2.

The Organizational Context

Firms typically contemplate adopting a more coordinated GMS, once they have significant presence in several countries and regions. Since local markets will never be exactly the same, a proposed global strategy will generally not be welcomed by the country managers. The existing local operations will have to be convinced to adopt the new global strategy. Thus, a GMS is always top-down, not bottom-up, and it is easy for antiglobalization sentiments to stir even within a multinational company.

The typical solution to this problem is to allow country managers to be involved in the formulation of the GMS, and to form cross-national teams to participate in the implementation. It is also common to designate one country the "lead" market for the strategy, and use its current strategy as a starting point for the global strategy. This lead country is typically one of the larger markets and one where the firm has a strong market share. In multibrand firms, it is also common to limit a global strategy to one or two brands, allowing the local subsidiaries to keep control of some of their own brands.

Global Segmentation and Positioning

The firms most likely to engage in GMSs are those present in global markets. Global markets are those where customer needs, wants,

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 2 General pros and cons of global marketing strategies.

	Pros	*Cons*
• Revenue side	Reinforced message, unique idea Spillover of brand awareness Enhanced liking (mere exposure)	Culturally insensitive Antiglobal target Vulnerable to gray trade
• Cost side	Reduces duplication, waste Uniform product design, packaging, advertising Quantity discounts in media buy	Requires managerial time Lowers morale in subsidiaries, agencies

and preferences are quite similar across the globe (*see* MARKET DEFINITION). Typical product categories are technology products, including consumer electronics, cameras and computers, branded luxury products, and also apparel, personal care, and entertainment categories where, for certain segments, globally standardized products are desired by all. By contrast, in multidomestic markets such as food and drink, where preferences are more culturally determined, global coordination is less common (*see* CUSTOMER ANALYSIS). For example, ACNielsen's cross-national data suggest there are only 43 global brands in the consumer packaged goods categories found in the typical supermarket (ACNielsen, 2001).

Global segmentation. The need to target similar segments in different countries is an attempt to minimize the drawbacks of a coordinated global strategy (*see* MARKET SEGMENTATION AND TARGETING). A typical cross-national segment targeted with a standardized product is the teenage and young adult segments, where preferences are allegedly very similar even for food and drink categories. Coca Cola uses the same one-word slogan "Always" around the world. Nike is positioned with a rebellious image in many countries, even though the particular sports associated with Nike differ by country. Technology brands such as the iPod have usually even more coordinated global strategies, with synchronized rollouts of new models across countries.

Global marketers might use a two-stage approach to market segmentation (*see* MARKET SEGMENTATION AND TARGETING), first grouping countries into similar regions to increase the chances of finding homogeneous subgroups within each region. Often the first step amounts to selecting a trade bloc, such as the European Union. As research has documented, many global strategies are, in fact, more regional than global (Rugman, 2005).

A GMS can also be successful if the firm has managed to change local preferences. A new product entering a local market will usually change preferences to some degree, whether by new features, promotion, or price. This is the basis for the extreme standardization proposed by Levitt in his seminal 1983 HBR (Harvard Business Review) article, where he suggests that "everybody" likes the same products. Examples of this abound. IKEA, the Swedish furniture retailer, has changed the market for furniture in many countries – it uses a very standardized and coordinated marketing strategy, focusing around its simple and functional furniture, annual catalog, and warehouse stores. Starbucks, the American coffee chain, also has re-created and enlarged a mature market in several countries with its new coffee choices, novel store layouts, and wider menu. In other cases, changes in the environment have affected preferences so as to make standardization possible. "Green" products are naturally targeting global segments, as are the lighter beers, the bottled waters, and the shift to wines. Such global segments naturally induce companies to adopt GMSs.

Global positioning. The main issue in global positioning (*see* POSITIONING ANALYSIS AND STRATEGIES) is whether the product offering should be positioned the same way everywhere or not. Complicating the issue is the fact that even with complete uniformity of the marketing mix,

the arrived-at position may still differ between countries. A classic example are Levi's jeans, whose rugged outdoors image places it in a mainstream American lifestyle segment, but becomes a stylish icon in other countries. Also, as this example illustrates, even if a brand wants to be seen as "global," its position is typically affected positively or negatively by its country of origin.

A fundamental factor affecting transferability of a position is the actual use of the product. A food product such as apples might be consumed as a healthy snack in the West ("An apple a day keeps the doctor away" as the saying goes). But in Japan, apples are a favorite item in the gift-giving season, placing a premium on color, packaging, and price – hardly the same positioning.

Even without such dramatic usage differences, differences in economic development and cultural distance, in general, are main factors influencing the potential for an identical position. A Ford car may be positioned as a functional value product in Europe, but might be a status symbol in a poor country. First-time buyers in emerging markets rarely view products the same way as buyers in the more mature markets, where preferences are well established. For example, the successful Buicks offered to new customers in China offer quite different benefits from those offered Buick customers in the United States, even though the product is largely the same.

The strength of local competition (*see* COMPETITIVE ANALYSIS) is also likely to vary across countries, affecting the positioning. Where domestic competitors are strong, a foreign brand that is a mainstream brand at home will typically attempt to target a niche abroad. This applies to many European brands including Heineken, Illycaffe, and Volvo. In other cases, a company with a niche position at home may target a more mainstream position in another market – an example is Japanese Honda in the US auto market. In global markets, where often the same global players compete in the major foreign markets, positioning is more likely to remain constant across the mature markets. Examples include automobiles, with the global players occupying very similar positions in most markets. This is less true for new product categories that are still in the growth stage in many countries and the brands are not equally well known everywhere. Cell phone makers Nokia, Samsung, and Sony-Ericsson occupy quite different positions in each market.

The stage of the life cycle (*see* STAGES OF THE PRODUCT LIFE CYCLE) is also likely to vary across countries, affecting how well a particular position can be transferred. In the early stages, with preferences still in flux, a strategy based on the positioning in a lead country may not be very effective in a new country. Thus, the first automatic single-lens reflex camera was introduced by Canon as a mainstream product in Japan, but a specialty product for more professional photography overseas. In emerging countries with their pent-up demand, however, even new consumers aspire for the best products in the leading markets. This is why some Western companies (such as Electrolux, the home appliance manufacturer) will position themselves at the top of the market even in a country like Russia.

The typical strategic assumption is that a globally uniform positioning requires similarity of culture, of competition, and of life cycle stage. However, even in countries where one or more of these requirements are not met, a standardized global positioning may still work. For example, when global communications have made the brand name already well known, a global strategy may work even in a multidomestic market. McDonald's successful entry into many emerging markets is a case in point. And even where domestic competition is strong and would suggest a niche positioning, external events may shift the market in favor of a newcomer. This happened, for example, when the Japanese autos entered the American market and gained strength during the 1970s oil crisis. But these are exceptions and are certainly not automatic, as Coca Cola learned in India when local ThumsUp rebounded (see the section Global Brands).

The Global Marketing Mix

Global products and services. Standardization of the product or service is usually a major feature of a global MARKETING MIX. "Product Standardization" means uniformity of product or service features, design, and styling. There are several advantages to such standardization, including those listed in Table 3.

Table 3 Advantages of product standardization.

- Cost reduction
- Improved quality
- Enhanced customer preference

Table 4 Disadvantages of product standardization.

- Off-target
- Lack of uniqueness

The advantages are mainly on the cost side – scale economies from the larger number of identical units produced. But there are also quality advantages involved. With longer series, there is more reason to invest in specialized technology, machine tools, components, and parts, yielding higher and more consistent quality. Finally, there is a possible positive demand effect on customers. Because of the prevalence of the products and designs, the "mere exposure" of individuals to the products engenders a positive impact on preferences. This is an effect which partly depends on competitive imitation – when most cell phones feature a built-in camera, consumers "want" a camera with their cell phone (*see* COMPETITIVE ANALYSIS).

The disadvantages of product standardization are mainly on the demand side (see Table 4).

Apart from the case of pent-up demand in an emerging country, standardized products rarely manage to target precisely a specific segment in a new country market. They are at least slightly off target. This is not always such an obstacle to success. First, preferences may change – the standardized product may offer features not offered before in that market. Honda's 1970s entry into the US car market exemplifies this case, with the car offering both fuel efficiency and sportiness. Second, a mispositioning may be overcome by a strong brand name. The McDonald's entries in emerging markets fall in this category. Third, the entering product may well be sold at a low price – its scale advantages can allow such a strategy. This was the strategy followed by Samsung before its later drive toward a strong global brand (Quelch and Harrington, 2004).

Since the typical multinational company does manufacturing in a large number of country subsidiaries, the need for scale has sometimes made it necessary to designate local production sites as suppliers for the whole world. Toyota's Kentucky plant produces the Camry for global distribution. The BMW Z4 sports car is only produced in South Carolina. Apple computers are all produced in Taiwan. In general, however, the risks of local strikes and political conflict make most companies assign production to more than one site.

From a marketing perspective, a uniform product or service is often less acceptable locally. Of course, some localization is always necessary in any case – electric appliances face different voltages and plugs, safety regulations differ between countries, and homologation requirements differ. But the more critical issues revolve around customer acceptance. What is seen as a good product or service in one market might not be acceptable elsewhere.

The fact is that there are relatively few products and services that are identical around the world. One would expect that products in global markets, such as technology products, would be identical. But generally speaking, PCs and cell phones are smaller in Asia, automobiles have a harder suspension in Europe than in the United States, and even stereo speakers vary slightly in bass level between North America (heavy on bass) and Asia (where smaller apartments places the listener closer). The classic failure by Euro Disney to transfer its American theme park unchanged to France is a good example of misguided standardization of a service product. Despite the success of the strategy in Tokyo, the Euro effort fell flat for many reasons, one of which was the no-alcohol rule inimical to Continental Europeans.

Luxury products are usually the same across the globe, and utilitarian items such as automobile tires, toothpaste, and kitchen utensils can be standardized. But products such as shampoos, soaps, and personal-care items need to take account of hair types, skin color, and water quality to perform satisfactorily. Coca Cola's level of sweetness differs across countries, McDonald's menu is adapted to country

preferences (partly to reduce antiglobalization protests), and apparel manufacturers have to make adjustments for different body proportions between Western and Asian peoples.

To deal with these adaptations while trying to retain some scale economies, companies resort to two solutions. One solution is to use the same basic design or "platform" for the product, and then adapt by adding alternative features at the later stage of manufacturing. This is common in automobiles, where the platform involves the chassis on which the body is then fitted. But the concept is also used in the manufacturing of electronic products, computers, and home appliances. This is the solution adopted by Coca Cola and McDonald's as well.

A second related option is to break up the product into component modules that can be produced in large series to gain the scale advantages, and then produce different products by different combinations of modules. This has become a very prominent manufacturing strategy for large companies, since it allows the different modules to be outsourced and offshored. The manufacturing process then becomes a simple assembly process, which can then be done locally, if necessary, to gain lower tariff rates. This allows the company to "mix and match" features for different country markets, which helps adaptation to local preferences. It also helps to make the products in different markets somewhat different, helping to limit gray trade (more on gray trade below).

In the end, companies do not need to offer identical products everywhere in order to gain the scale economies of product standardization. Thus, a company can develop a coordinated global strategy even without a completely standardized product. But it is almost impossible to develop a GMS without a strong global brand.

Global brands. Keeping the same brand name everywhere has become the signature feature of a global marketer, and "global branding" has become an obsession among many multinationals. For example, the Interbrew (now InBev) company's analysis of the Heineken advantage in profitability draws the conclusion that it is the lack of a global brand that depresses its own bottom-line performance. Hence, top management has decreed that Stella Artois be Interbrew's global "flagship" brand (Beamish and Goerzen, 2000).

Three definitions follow:

- *Global brands* are brands that are well known and recognized in all major markets of the world. (e.g., Sony, Mercedes-Benz, Microsoft, Nokia).
- *Regional brands* are brands that are the same across a region (e.g., P&G's Ariel in Europe is Tide elsewhere; Acura is Honda Legend in Asia).
- *Local brands* are brands found in only one or two markets (e.g., Suntory whisky in Japan, A&W root beer in the United States, and the Trabant car in former East Germany).

Strictly speaking, the brand may be global although the product is not available everywhere – as happens to be the case for Rolls Royce as well as for Coca Cola. This usually means there may be a pent-up demand for some global brands, as was seen when McDonald's entered Russia in the 1990s.

Global brands have received increased attention from top management in many multinationals because of the importance of brand equity as a financial asset (*see* PERCEPTION OF BRAND EQUITY). Expanding into new markets is an obvious way of building further financial equity, which is usually calculated by simply aggregating projected revenues across country markets. Not surprisingly, most top brands in terms of financial equity are global. But a strong brand not only needs reach across countries, it also needs allegiance from local customers. As global brands have stretched further to build financial equity, local brands have been able to defend their turf by staying closer to their customer and building affinity, or what may be called *soft equity* (*see* CUSTOMER EQUITY).

Recognizing this, many global companies not only market their global brand in a country market but might also buy up a successful local brand and retain its brand name – and customers. One example is Coca Cola in India. After its reentry in 1993 (Coca Cola had exited India in 1977 instead of giving up its secret formula), Coca Cola acquired ThumsUp, a leading local producer, with the idea of replacing its brand

with Coca Cola. But after several efforts at withdrawing ThumsUp and launching its own brand, Coca Cola finally gave up and shifted marketing resources to ThumsUp. The problem was that Coca Cola was positioning itself as the young teenage drink, just as it had done in many countries, while in India, where alcohol consumption is very limited for religious and cultural reasons, the main cola market was among young adults who in other parts of the world were drinking beer. ThumsUp in India is a rebel's drink, hardly the image of Coca Cola.

The most clear-cut advantages of global brands are the cost efficiencies from scale and scope. The typical benefits to global brands are several (see Table 5).

The cost efficiencies tend to come from the ability to produce identical products and packaging in long series, and also because global brands can draw on uniform global promotions (more on this below). Demand spillover is a result of the increased exposure to the same brand in many places, especially useful when customers are global. The growth of international tourism has been a strong driver of global brands. The status and esteem advantages have been shown by researchers, especially prominent in less-developed countries. While some research has demonstrated a high quality perception for global brands, the more firmly established finding is that global brands tend to have a more consistent quality than local brands.

The disadvantages of global brands become advantages for local brands. Local brands can usually count on the advantages in Table 6.

Of course, none of these advantages come without effort and disciplined application by the firms, whether global or local. The arrival of global brands into many markets has been a challenge for many local brands who think that local consumers will automatically stay loyal.

Table 5 Benefits to global brands.

- Scale and scope economies
- Demand spillover
- Global customers
- High esteem, status
- Consistent quality

Table 6 Advantages for local brands.

- Local brand affinity
- Motivated local employees
- Prodomestic (and Antiglobalization) sentiment

Global brands, as Naomi Klein claims, have changed the playing field – but they have not, as she claims, simply dominated local brands (Klein, 2002).

Global pricing and distribution. In GMSs, pricing, and distribution are more closely connected than at home. The reason is not that the costs involved in distribution (transportation but also insurance and custom duties) necessarily raise the final price to the customer. Such straight "price escalation" does not usually occur except in one-time transactions. Many multinationals have strong home market "cash cows," and when faced with more intense competition in foreign markets they reduce prices by lowering transfer prices to their subsidiaries. Some firms also use foreign markets as an easy way out of overcapacity, applying marginal cost pricing procedures (although these can run afoul of dumping laws). And the improved efficiency of global transportation, thanks to global express carriers and consolidated shipment procedures, means that geographic distance is no longer the trade barrier it once was. Transportation costs are typically a small proportion of the total price paid (*see* MARKETING CHANNEL STRATEGY).

The strong connection between pricing and distribution rests more directly on another phenomenon. The ease of transportation, coupled with differing local prices and currency fluctuations, are what provide the margin that allows for arbitrage opportunities for customers to buy branded products cheaper abroad. This is an instance of so-called "gray trade" – the importation of branded products through other than authorized channels (*see* MULTICHANNEL MARKETING). It is the rise of gray trade that force multinationals to decide pricing and distribution strategies jointly – and even multinationals that would otherwise not contemplate a global strategy, have to find a way to align prices to avoid such trade.

Gray trade affects a number of multinationals. For example, there are numerous stories of Asian contract manufacturers who make branded products for Western multinationals on day shifts, and then produce an added batch of identical products on the night shift. These products are then shipped abroad at low costs, distributed via indirect channels perhaps from a third country, and finally appear on the various markets in the West. In other cases, gray trade involves Western distributors (large European retailers, for example) who acquire goods in a low-priced country and then sell it at higher prices at home. Britain's Tesco chain sent buyers to Wal-Mart stores in the United States to buy Levi's for resale at home, a practice that was stopped by a European court. Another popular form of gray trade, "shopping tourism," is what happens when overseas trips are arranged for tourists to buy products cheaper in a foreign country.

The drivers of gray trade include the factors in Table 7.

Gray trade is not usually illegal. By contrast, trade in counterfeits (fake products) is illegal and vigorously opposed by multinationals that fear the loss of revenue and dilution of their brand name. But there are similar negative effects from gray trade as from counterfeits. Gray trade strains the relationship with authorized channel members since channel members face intrabrand competition. There may be legal liabilities, usually involving warranties that cannot be honored. There is also a risk of erosion of brand equity because of the lower price in the market. Seiko, the Japanese watch maker, has failed to establish itself as a strong premium brand partly for this reason. And gray trade complicates global coordination when one country realizes a sudden influx of gray goods. Even though companies cannot take legal action, when faced with gray trade, they have to engage in relationship building with distributors, screen orders carefully, and monitor shipments.

Table 7 Selected drivers of "gray trade."

- Transportation is global and efficient
- Trade barriers are low
- Products and brands are standardized
- Communication is global

Because of pricing regulations, the multinational producer cannot usually dictate retail prices in local markets (*see* PRICING STRATEGY). Nevertheless, to achieve the desired brand position, they can use suggested retail prices. These are the prices that have to be coordinated across countries because customers can purchase the products anywhere in the world. But to coordinate prices is difficult, for several reasons as shown in Table 8.

One single global price is unrealistic. Even though a business-to-business company such as Boeing, the aircraft builder, quotes all its prices in American dollars, exchange rate problems still arise for its customers. For many global companies, the solution is to devise "pricing corridors," centered around a desired positioning price. The "corridors" are the limits of prices between which the local price may vary without interference from headquarters. The price corridors should reflect not only demand and competitive pressure in the local market but also the differences in exchange rates and likelihood of gray distribution – a very difficult balancing act. In addition to formal corridors with a centralized positioning price, the global pricing coordination typically involves informal coordination with the local subsidiary to allow flexibility (Assmus and Wiese, 1995).

Global marketing communications. Next to global brands, the most visible aspect of a GMS is perhaps global advertising. *Global advertising* can be defined as media advertising that is more or less uniform across many countries, often, but not necessarily, in media vehicles with global reach. Although global appeals had been used previously in promotions – IBM's

Table 8 Why global coordination of prices is difficult.

- Currency exchange rates fluctuate
- Local distributors are independent
- Import prices to subsidiaries have to consider tariffs, taxes.
- Local competition varies across countries

Table 9 Major drivers of global advertising.

Supply Side
• Global ad agencies
• Global media
Demand Side
• Global customers
• Preference convergence

global "Think" slogan appeared as early as the 1920s – global advertising arrived with the advertising agency Saatchi & Saatchi's television commercial "Manhattan Landing" for British Airways in the early 1980s. With increasing globalization and the stress on global brands, the momentum behind global advertising has been sustained despite antiglobalization and prolocalization sentiments around the globe. One contributing factor has been the rise of the Internet and the availability of many commercials on sites such as YouTube, where even local advertisement campaigns potentially have global reach.

There are several forces behind the need for integrated global communications (*see* INTEGRATED MARKETING COMMUNICATION STRATEGY). One can distinguish between supply-side drivers and demand-side drivers, as shown in Table 9.

On the supply side, the emergence of consolidated global advertisement agencies has played a significant role in generating more global advertising. Although in many ways, the agency globalization has been a response to the globalization of the client firms – global managers find it useful to deal with the same agency in different parts of the world – once established, the global agency will naturally want to leverage its global capabilities (as in the Saatchi & Saatchi case). The global agency can also claim superior production values with a global campaign, since more resources can be used for one television commercial that is going to be shown around the world. The emergence of global media – the BBC, the Sky channel, the CNN, Financial Times, and so on – and their consolidation into global media companies such as Viacom, Bertelsmann, and Time Warner has also encouraged the development of global advertising and promotions that can be used effectively everywhere. There are cost savings in buying all media from one consolidated source.

On the demand side, customers are increasingly global. Consumers now travel much more than before as the lower cost of travel have made for many more tourists, and for business-to-business products, the customers are often multinational companies. In addition, with global communication producing spillovers between countries, local preferences change and allow the penetration of global brands in local markets using standardized appeals. Globally coordinated advertising thus becomes a natural complement to the global brand. It has given us the "Always" of Coca Cola, the "Our Passion" of Microsoft, the "Do you dream Sony?" and "The Ultimate Driving Machine" of BMW. Global Internet ads are now common, with companies such as IBM and Heineken producing commercials on their Web home pages and repeated on the YouTube.

In the general case, however, not all the marketing communications of a global company is globally coordinated. First of all, media advertising is only one of several promotional tools. Many retail promotions such as POP (point of purchase), coupons, and free samples, are necessarily more localized. Regulations vary across countries. Not all countries' retail regulations allow contests, for example, and in many cases, coupon redemptions are denied by stores. Even naturally global promotions such as sponsorships of the Olympics and the World Cup often require local input to leverage the promotion effectively. In one notorious instance, Budweiser's sponsorship of the World Cup in Germany in 2006 encountered opposition from German brewers and consumers. In an effort to appease German drinkers, Budweiser made a local deal to allow the German-made Bitburger beer to be sold in stadiums, albeit in unmarked cups.

But even media advertising is rarely fully globalized. The motivation for local subsidiaries and their agencies to do their best creative work is enhanced with more autonomy. Not all media are equally available in all countries, and the costs vary considerably. Effectiveness also varies. In poorer countries, print media are usually less effective. By contrast, Europe on the whole

gives much greater weight to print than other countries. The Internet, a new and naturally global channel of communications, has still not penetrated all corners of the world.

In addition, the advertising message often has to be adapted. Linguistic, cultural, and religious differences can prevent standardization of advertising messages and render symbols inappropriate. IKEA's use of the Moose, successful in Europe and Canada, was too provincial in the United States, where a simpler blue-and-yellow logo is used. Product usage may not be the same, making a uniform appeal miss the target. Ice cream is bought for its nutritional value in poorer countries, not so in advanced markets. This means the local subsidiary and its advertising agency usually have to do more than merely translate a message. Not all local adaptations work of course. A Nike advertisement with LeBron James, the American athlete, slaying a dragon turned out to be a failed local adaptation in China, where dragons are a force for good.

Because of these issues, most companies dedicate only a portion of their total advertising budget to a global campaign. An example comes from Samsung, the Korean electronics maker who has moved ahead of Sony in global brand value in the last few years (*see* BRAND VALUE). In 2003 about 40% of above-the-line budgeting was for global advertising. The remaining 60% of a $1 billion budget was allocated to local and regional subsidiaries for their nonglobal spending on specific products and markets. While the localized campaigns may employ local agencies, it is common for the global campaign to be handled by one large global agency, FCB (Foote, Cone, and Belding), in the Samsung case.

Most global companies place strict limits on how their name should be portrayed, including fonts and coloring. Sony requires all uses of their name and logo to be approved at headquarters in Tokyo before release. There are a few identical ads used, particularly in print advertising. The Marlboro cowboy can be seen around the world. The well known Absolut advertisements, with the shape of the bottle contoured in the advertisement, are used in many countries, sometimes localized. These are examples of the notion that a great idea can be used everywhere, a common assertion in advertising circles.

But even in a global campaign there is usually some variety. The typical form of global advertising in television follows what is known as pattern standardization. Here the advertisement visualization is adapted to local culture and language, with recognizable local spokespersons and actors, and a story that has local appeal. The brand name and logo are identical, and the final slogan is usually translated directly. Instead of actors speaking, voiceovers allow local language to be superimposed on a commercial. In this form, global advertising is becoming more common today, creating a unified image of corporations and brands as well as countries and places.

Summarizing, the advantages of global communications include those shown in Table 10. The potential disadvantages are also several, as shown in Table 11.

In the end, most companies play it safe, with some global uniformity but allocating the majority of the funds for regional and local adaptation of communications.

CONCLUSION

GMSs have become increasingly important with the internationalization of business and

Table 10 Advantages of integrated global communications.

- Consistency of brand communications
- Media spillover
- Cost savings
- Improved production
- Leveraging a great idea

Table 11 The disadvantages of integrated global communications.

- Images and symbols might not be locally acceptable
- Appropriate media might not be available
- Product usage is not the same
- Local creativity can be stifled

globalization of markets. Even though they are characterized by centralized coordination and streamlining to achieve scale and scope economies, localization, and adaptation are becoming increasingly important as emerging markets rapidly manifest culturally and ethnically differentiated consumer demand. Global success depends crucially on striking the right balance between uniformity and local adaptation. This balance, as we have seen, involves both top-down leadership and sensitivity to local markets – a true managerial challenge.

Bibliography

ACNielsen (2001) *Reaching the Billion-Dollar Mark: A Review of Today's Global Brands*. ACNielsen Inc., Chicago.

Alden, D.L., Steenkamp, J.-B.E.M., and Batra, R. (1999) Brand positioning through advertising in Asia, North America and Europe: the role of global consumer culture. *Journal of Marketing*, **63** (1), 75–87.

Assmus, G. and Wiese, C. (1995) How to address the gray market using price coordination. *Sloan Management Review*, **36** (3), 31–42.

Beamish, P.W. and Goerzen, A. (2000) The Global Branding of Stella Artois, Case 900A19. University of Western Ontario, Ivey School of Business, p. 23.

Dawar, N. and Parker, P. (1994) Marketing universals: consumers' use of brand name, price, physical appearance, and retailer reputation as signals of product quality. *Journal of Marketing*, **58** (2), 81–95.

Ghemawat, P. (2007) Why the World isn't flat. *Foreign Policy*, **159** (5), 60.

Holt, D.B., Quelch, J.A., and Taylor, E.L. (2004) How global brands compete. *Harvard Business Review*, **82** (9), 68–81.

Klein, N. (2002) *No Logo: No Space, No Choice, No Jobs*, Picador, New York.

Quelch, J. and Harrington, A. (2004) *Samsung Electronics Company: Global Marketing Operations*. Case 9-504-051. Harvard Business School, Boston, MA, p. 32.

Roth, M.S. (1995) The effects of culture and socioeconomics on the performance of global brand image strategies. *Journal of Marketing Research*, **32** (2), 163–175.

Rugman, A. (2005) *The Regional Multinationals*, Cambridge University Press, Cambridge.

Zou, S. and Cavusgil, T.S. (2002) The GMS: a broad conceptualization of global marketing strategy and its effect on firm performance. *Journal of Marketing*, **66** (4), 40–56.

go-to-market strategy

Raji Srinivasan

The firm's go-to-market strategy involves the development of a marketing program to create, capture, and sustain value with the firm's offering for the firm's customers, given that the firm has identified a product to market.

Typically, the firm's go-to-market strategy can be conceptualized into two stages: (i) market segmentation, target market selection and positioning, given the product's (good or service) unique selling proposition, (ii) the marketing mix for the product, given the market segment the firm seeks to target for the product. These two stages in the firm's go-to-market strategy are discussed below.

Market Segmentation, Target Market Selection, and Positioning

The centerpiece of the firm's go-to-market strategy is the customer. The firm must identify the customers it seeks to address with the product on hand. The appropriate identification of the target market is a prerequisite for the development of a successful go-to-market strategy, allowing the firm to focus its marketing efforts on the appropriate customer segments.

Market segmentation calls for dividing the market into groups of customers, market segments, each with distinct characteristics (e.g., demographics, psychographics, behaviors, lifestyles). The objective of market segmentation is to develop groups of customers who substantially differ from each other, but are homogenous within the group (Rao and Steckel, 1998). Segmentation requires the following three steps from the marketer: (i) they understand customers' needs and the benefits that customers seek from the product, (ii) segment the market and develop prototypical customer profiles either based on customer demographics and/or benefits sought, and (iii) identify the observable variables (e.g., demographic characteristics) most likely to be discriminatory among the different customer segments.

Target market selection involves evaluating each market segment's attractiveness and the identification of one or more segments to enter. In addition to the identification of the market segment, the marketer must also have insights on their firm's strengths (and weaknesses) and that of their competitors. The objective of target market selection is to identify target markets that are most profitable, given the product's features and benefits, costs of entering the market, and the market segment's characteristics. The crucial step in target market selection is a differential advantage analysis, an understanding of the firm's offerings, relative to the competitors' offerings. In addition to an understanding of competitors' offerings, it is also useful to anticipate competitive reactions so that such reactions can be incorporated into the development of the marketing program for the product.

Positioning is defined as the marketer's effort to identify a unique selling proposition for the product, so that the product occupies a clear, distinctive, and attractive position relative to competing products in the minds of target customers (Ries and Trout, 1981). A good positioning statement should address the following three questions: Who are the customers? What is the set of needs that the product fulfills? Why is the product the best option to satisfy those needs? A good positioning statement is aimed at potential customers, reflects clear differentiation of the product from competitive offerings, and guides the development of the marketing plan for the product.

Marketing Mix

The firm's *marketing mix* describes the set of activities comprising a firm's marketing program (Borden, 1964). Borden specified twelve marketing mix elements: product planning/merchandising, pricing, brandings, distribution channel, personal selling, advertising, promotions, packaging, display, servicing, physical handling, and market research. Over time, these 12 elements have been considered as the four Ps – product, place, promotion, and price.

Product decisions involve aspects of the product which include the physical features of the product, brand name, and financing plans (Crawford, 1997). A broad conceptualization of product is the key to achieve effective differentiation from competitors' offerings. Product decisions also involve considerations

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

of product line length (how many products to cover different market segments?) and product line depth (how many types of a given product?)

Place decisions refer to decisions about the firm's distribution channels or how the firm "goes to market." A customer-behavior-driven approach to channel design is advocated to ensure that customers' needs are met in the most cost-efficient manner (Stern, El-Ansary, and Coughlan, 1996). An important issue with respect to channel design is that the firm should not only decide which types of channels to employ (distributors, retailers, jobbers) but also determine the functions and the appropriate channel incentive design for the different channel members. A key aspect of channel design is the identification of policies and procedures that will be used to ensure that the various members in the channel perform their designated functions.

Promotion decisions refer to the set of ways in which the firm communicates with its customers with a view to increase awareness of the product, and persuading customers to either try the product or repeat purchase it. Effective marketing for a product requires an integrated marketing communications plan combining personal selling efforts and impersonal approaches including advertising, sales promotion, and public relations (Aaker, Batra, and Myers, 1992). A key to developing an effective communications program is to understand the people involved in the decision-making process, and their motivations and needs. In today's environment with the World Wide Web playing an important role in consumers' purchase decisions, the role of user-generated content including consumer blogs and ratings are proving to be as influential in consumer decision making as firm-provided information.

Price decisions refer to decisions about the product's pricing. Typically, when the product, place, and promotion are decided upon, the consumer's perception of value of the product is determined, so that the price range of the product may be in place. There may be several considerations that guide pricing decisions such as the cost of the product, which includes manufacturing, promotion, and distribution costs, cost structure, competitors' pricing in the market, consumers' price elasticities, the availability of substitutes for the product, the firm's profit objectives, the firm's market objectives, and so on (Dolan and Simon, 1996).

An effective go-to-market strategy is both an art and a science that involves the integration of customer insights and analytics to ensure that there is a match between the firm's target market and the marketing mix developed for the product so that both consumer needs and the firm's performance objectives are met.

See also *first-mover (pioneer) advantage; later mover (nonpioneer) advantage; marketing strategy*

Bibliography

Aaker, D.A., Batra, R. and Myers, J.G. (1992) *Advertising Management*, 4th edn, Prentice-Hall, Englewood-Cliffs.

Borden, N. (1964) The Concept of the Marketing Mix, in *Science in Marketing* (ed. George Schwartz). John Wiley & Sons, Inc. New York.

Crawford, M.C. (1997) *New Product Management*, Irwin, Homewood.

Dolan, R.J. and Simon, H. (1996) *Power Pricing*, 5th edn, Free Press, New York.

Rao, V.R. and Steckel, J.H. (1998) Segmenting markets: who are the potential buyers, *Analysis for Strategic Marketing*, Addison-Wesley, Reading.

Ries, A. and Trout, J. (1981) *Positioning, The Battle for your Mind*, McGraw-Hill Inc., New York.

Stern, L.W., El-Ansary, A.I., and Coughlan Anne, T. (1996) *Marketing Channels*, 5th edn, Prentice-Hall, Englewood-Cliffs.

innovation diffusion

Vijay Mahajan

The diffusion of innovation is defined as the process by which that innovation is communicated through certain channels over time among the members of a social system (Rogers, 1995). The diffusion process consists of four key elements: innovation, communication channels, time, and the social system.

As a theory of communication, diffusion theory's main focus is on communication channels, which are means by which information about an innovation is transmitted to or within the social system. It consists of both the mass media and interpersonal communications. Members of a social system have different propensities for relying on mass media or interpersonal channels when seeking information about an innovation. Interpersonal communications, including nonverbal observations, are important influences in determining the speed and shape of the diffusion process in a social system.

Recent literature suggests two additional mechanisms for innovation diffusion: signals and network externalities (Peres, Mahajan and Muller, 2009). Signals are defined as any market information other than personal recommendation that can be used by a potential adopter to make an adoption decision (e.g., the number of previous adopters may signal better product quality). Network externality refers to the observation that the utility of some products or services may increase as more consumers adopt the new product (e.g., Internet or mobile phones).

Consumer behavior researchers in marketing have focused on developing and evaluating hypotheses related to the adoption and diffusion process. For example, how do the innovation attributes (e.g., complexity) affect the adoption/diffusion process? What are the differences between the various adopter categories (e.g., innovators vs laggards)? How does the negative and positive word-of-mouth affect the diffusion speed? What is the role of opinion leaders in the diffusion process?

The marketing management literature has focused on the implications of the above hypotheses for targeting and developing marketing strategies aimed at potential adopters. For example, how should a new product be positioned for innovators versus laggards? When might it be possible to skip the innovators and directly go to the mass market? What is the relationship between sampling and the diffusion speed?

Researchers in marketing science have contributed to the development of diffusion theory by suggesting analytical models that describe and forecast the diffusion of an innovation in a social system. Among these models, the best known is the Bass model (Bass, 1969). This model assumes that the potential adopters comprise of two groups. The first group is influenced only by the mass media-communication and is termed as "innovators." The second group is influenced only by the word-of-mouth communication and is termed as "imitators." Since its publication, a number of applications, estimation procedures, and extensions have been reported incorporating the effect of marketing mix variables, competition to study the diffusion of the various brands and their influence on each other's growth, and diffusion over markets and countries (spatial diffusion) (Mahajan, Muller and Wind, 2000). The Bass model yields a unimodal (single peak) curve for the incremental adoptions over time and an S-shaped curve for the cumulative adoptions over time. The recent studies suggest that the Bass curve should include three other elements or turning points: take-off, saddle, and technological substitution.

The Bass model starts with the adoption of an initial group of adopters, yet does not provide explanations as to the mechanisms that led to this adoption. Following take-off, the Bass model predicts a monotonic increase in sales. However, in some markets, this increase might be nonmonotonic, and a sudden decrease in adoption may occur after an initial rise (referred to as *chasm* by Moore) and then increase again producing a saddle in the early growth of the innovation. The Bass model also assumes that diffusion is terminated by the saturation of the market potential. However, in practice, new products are substituted with more advanced products and technological generations.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Bibliography

Bass, F.M. (1969) A new product growth model for consumer durables. *Management Science*, **15** (1), 215–227.

Mahajan, V., Muller, E. and Wind, J. (2000) *New Product Diffusion Models*, Kluwer Academic Publishers, New York.

Peres, R., Mahajan, V. and Muller, E. (2009) Innovation diffusion and new product growth models: a critical review and research directions, Working Paper.

Rogers, E.M. (1995) *Diffusion of Innovations*, Free Press, New York.

integrated marketing communication strategy

Ricardo Villarreal

IMC Background

In the early 1990s, marketing academics and professionals began to understand the importance of marketing communications consistency in achieving marketing goals and objectives. This understanding was formally organized as integrated marketing communications (IMC) (Schultz, 1991; Moore and Thorson, 1996). Overall, IMC plays an important role in a marketing strategy, especially in terms of branding (*see* BRAND VALUE).

Despite its more than 20-year history, the concept of IMC is still contentious. One fundamental issue with the IMC is how to conceptualize and therefore define it. The conceptual/definitional problem of IMC stems from the fact that it is both a concept and a process (Kliatchko, 2005). A number of conceptualizations and definitions have appeared since the 1980s (Shimp, 2007); for a review of definitions see Kliatchko, 2005. A second fundamental IMC issue is how to measure its impact.

IMC and The Environment Within Which it Developed

IMC emerged as a natural evolution in marketing communications, brought about by drastic changes in the marketplace (*see* MARKET/INDUSTRY STRUCTURE), media and communications, and consumer characteristics (Kliatchko, 2005). Thorson and Moore (1996) and Sirgy and Rahtz (2007) describe specific changes that helped in the formalization of IMC. One identified change is the *decrease in message credibility*. As a result of message clutter, the ability for a single message to have much meaning was rendered almost impossible. A *decrease in the cost of database marketing* (*see* DATABASE MINING AND MARKETING) also helped in ushering the IMC concept. IMC thinking increased in conjunction with the declining cost of gathering, storing, and analyzing consumer data. This allowed firms to identify segment-related characteristics (*see* MARKET SEGMENTATION AND TARGETING) which lead to identifying important message-related characteristics. A third change in the marketplace was the *increase in the cost of and decrease in the effectiveness of mass media communications*. The number of people watching TV, listening to radio, and reading newspapers began to drop, while the cost to place messages in these media increased. In addition, other media, such as cable, online content, and computer gaming negatively influenced the effectiveness of traditional media. An *increase in mergers and acquisitions of marketing communications agencies* was another change that helped usher in the IMC concept. As a result, agencies were able to put IMC thinking into practice by providing IMC-related services and tasks. Another change was the *increase in media and audience fragmentation*. In the 1980s, the media industry experienced an explosion of content-specific channels and programming, such as sports channels (e.g., ESPN), educational channels (e.g., Discovery Channel), and gender-specific channels (e.g., Oxygen). This fragmentation was associated with an increasing need for multiple messages across a larger number of media and vehicles. The *availability and increase of parity and competing products* (*see* POINT OF DIFFERENCE AND PRODUCT DIFFERENTIATION) also influenced IMC thinking. Clutter at the product level lead to an increased need for product differentiation and the importance of branding (*see* BRAND STRATEGY). Finally, there was a *shift in information technology*. The growth and speed at which the technological environment was changing added to the number of possible new channels (*see* MARKETING CHANNEL STRATEGY) with which to interact with customers.

IMC Features

On the basis of a review of IMC definitions, Shimp (2007) identified five key features. The first feature states that the *consumer or business customer should be the starting point for all marketing communications activities* (*see* CUSTOMER ANALYSIS). A customer focus provides the basis for selecting the most appropriate messages and media. The second feature suggests *using any and all marketing*

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

communications tools that are up to the task. This suggests the importance of learning customers' lifestyles and media preferences in order to know the best context in which to reach them. The third key feature states that *multiple messages must speak with a single voice* in order to consistently and effectively communicate with customers. The fourth feature is *build relationships rather than engage in flings* (*see* CUSTOMER RELATIONSHIP MANAGEMENT). This implies using the tools that encourage repeat purchasing and enhance brand loyalty (*see* BRAND VALUE) whenever possible. The final feature states *do not lose focus of the ultimate objective: Affect behavior*. IMC's ultimate goal is to affect consumer behavior above and beyond intermediate goals such as creating brand awareness or attitude change.

IMC Tools

The difficulty in implementing IMC lies in selecting and managing possible IMC tools. There are generally eight media categories under which specific vehicle may be selected. Ultimately, brands have to choose those tools that are deemed appropriate for reaching targeted consumers (Shimp, 2007). *Media advertising* is any paid form of nonpersonal communication about an organization, product, service, or idea by an identified sponsor. Advertising tools include television, radio, magazines, and newspapers. *Direct response and interactive tools* (*see* DIRECT AND INTERACTIVE MARKETING) provide a system by which organizations communicate directly with target customers to generate a response and/or a transaction. Direct response tools include direct mail and telephone solicitation. Interactive marketing allows for a two-way flow of information, and tools primarily consist of online advertising and heavy use of the Internet. *Place advertising* generally refers to advertising seen outside of homes. Place advertising tools include billboards, bulletins, posters, transit, and cinema ads. *Store signage and point-of-purchase advertising* allows consumers to identify and locate businesses and can influence their store-choice decisions and impulse purchasing. Tools include external store signs, in-store shelf signs, shopping cart ads, and in-store radio and television. *Trade- and consumer-oriented promotions* provide extra value or incentive to the sales force, distributors, or the ultimate consumer. Tools include trade deals and buying allowances, display and advertising allowances, trade shows, cooperative advertising, samples, coupons, premiums, refunds and rebates, contests/sweepstakes, promotional games, bonus packs, and price-off deals. *Event marketing and sponsorship* allows brands to support events and causes that are congruent with the company/brand and consumer's values, interests, or positions. Tools include sporting events, arts, fairs, and festivals, as well as the sponsorship of particular causes. *Marketing-oriented public relations and publicity* are other communication options. Marketing-related public relations is a narrower aspect of general public relations in that it involves an organization's interactions with actual or prospective customers. Publicity is a nonpersonal communication about an organization, product, service, or idea that is not directly paid for nor run under identified sponsorship. *Personal selling* is direct person-to-person communication whereby a seller attempts to assist and/or persuade prospective buyers to purchase a company's product or service or act on an idea.

IMC Orientations

IMC can be categorized by either a tactical or a strategic orientation. Some believe the tactical orientation is the simplest and least effective, as it focuses primarily on executional factors such as message cohesiveness and continuity. On the other hand, the strategic orientation may be considered the next step in the evolution of understanding IMC. Its focus transcends the executional by integrating the larger corporate marketing strategy position.

Tactical. According to Parente (1996), *campaign continuity* means all messages communicated in different media through different marketing communications tools are interrelated. It involves making both the *physical* and *psychological* elements of a marketing communications campaign continuous. *Physical continuity* is the consistent use of creative elements in all marketing communications. This can be

achieved by using the same tagline, slogan, and trade characters across all advertisements and other forms of marketing communications. Stern (1996) states *psychological continuity* is a consistent attitude toward the firm and its brand(s). It is the consumers' perception of the company's "voice" and its "persona." Psychological continuity can be achieved by using a consistent theme, image, or tone in different ads and other forms of marketing communications.

Moore and Thorson (1996) identified four general IMC approaches that arguably are tactical in nature. The "*one-look*" *approach* is based on the strategy of creating a uniform look and feel for all executed messages from colors, fonts, visuals, and logos, in all forms of marketing communication tools. The *theme-lines or* "*matchbook*" *approach* attempts to coordinate all supporting forms of marketing communications around an advertising theme. The *supply-side planning approach* stems from a marketing communications firm's ability to package for the client a combination of marketing communications media, such as radio, TV, and print for a packaged price. The *ad hoc approach* may be found within integrated agencies or across agencies held by a common parent company.

Strategic Orientation. *Strategic orientation* is a systematic approach that guides firms to fully integrate strategies and objectives at the corporate and marketing organizational levels with all elements of the marketing communications campaign at the planning, execution, and monitoring stages (Sirgy and Rahtz, 2007). That is, every element in the campaign is designed to achieve specific strategic goals and media are selected with strategic goals in mind. The entire marketing and communications system of the firm is integrated toward delivering a unified strategic impact. Strategic marketing communications includes the implicit messages within execution of the four Ps, not just the communications aspect.

IMC Benefits

According to Kliatchko (2005), IMC benefits include "the reduction of media waste, improved coordination, centralization, and greater consistency of marketing communications programs, providing a positive effect on client budgets (*see* COMMUNICATIONS BUDGETING), and increased message impact and creativity . . ." (p. 12). IMC may reduce media waste by promoting only the selection of target-appropriate media. IMC may have helped lead to the restructuring of marketing communications within companies/agencies, as the conceptualization of IMC may have translated into a physical reorganization withincompanies/agencies. Presumably, reducing media waste and centralizing people and functions could have a positive budgetary savings effect.

IMC Challenges

Measurement. One of the most important IMC challenges is the inability to truly measure its effectiveness. It has been acknowledged that IMC effectiveness cannot be measured now (Schultz and Kitchen, 2000). IMC is difficult to measure because measurement attempts to focus on output as opposed to outcomes. In addition, "measurement problems are compounded by the fact that IMC programs consist of a variety of communication tools and measuring the interactive effects of all of these elements has proven to be extremely difficult."

Bilingual consumers. IMC is based on two assumptions. The first assumption is that the target consumer speaks only one language. The second is that if a given target consumer speaks more than one language then the consumer will not be exposed to messages for the same company/product in both languages. Research by Villarreal and Peterson (2008) and Villarreal, Blozis and Jeong (2008), however, suggest that these assumptions are not tenable for US Hispanics (*see* MARKET SEGMENTATION AND TARGETING), as adults and adolescents consume media in both Spanish and English. Villarreal and Peterson suggest that IMC thinking should be reconsidered to account for this phenomenon.

New technology and lack of control. New technology poses an issue for IMC in that this field is constantly and almost consistently changing. It may be problematic for manufactures to keep up with changes in new technology and integrate them (if they are relevant). In addition,

new media may become even harder to control in terms of message construction, positioning, and dissemination. For example, YouTube is fertile ground for spoof ads or videos criticizing a particular company. Such grass-root messages have the potential to spread widely and quickly on the Internet.

Conclusion

IMC is practiced and conceptualized differently by academics and practitioners alike. However, it is a concept with great appeal and theoretical and practical implications. IMC will most likely continue to evolve in conjunction with our understanding of and changes in our social, technological, and political environments. As Schultz and Kitchen (2000) noted "... IMC is still in a 'pre-paradigm state of development' and that its value will become more evident as further research and experience is obtained through the years."

Bibliography

Kliatchko, J. (2005) Towards a new definition of Integrated Marketing Communications (IMC). *International Journal of Advertising*, **24** (1), 7–34.

Moore, J. and Thorson, E. (1996) *Integrated Communications: Synergy of Persuasive Voices*, Lawrence Erlbaum Associates, Mahwah, NJ.

Parente, D. (1996) *Advertising Campaign Strategy: A Guide to Marketing Communications Plans*, The Dryden Press, Fort Worth, TX.

Schultz, D.E. (1991) Integrated Marketing Communications: The Status of Integrated Marketing Communications Programs in the US Today. *Journal of Promotion Management*, **1** (1), 99–104.

Schultz, D. and Kitchen, P. (2000) A response to 'Theoretical concept or management fashion?' *Journal of Advertising Research*, **40** (5), 17–21.

Shimp, T.A. (2007) Advertising promotion and other aspects of integrated marketing communications.

Sirgy, J.M. and Rahtz, D. (2007) *Strategic Marketing Communications: A Systems Approach to IMC*, Atomic Dog, Mason, OH.

Stern, B. (1996) Integrated communications: the company 'Voice' and the advertising persona, in *Integrated Communications: Synergy of Persuasive Voices* (eds J. Moore and E. Thorson), Lawrence Erlbaum Associates, Mahwah, NJ, pp. 87–101.

Villarreal, R., Blozis, S.A. and Jeong, J.J. (2008) Understanding adolescent Hispanic Consumers: advertising and branding implications. *Business Journal of Hispanic Research*, **2**, 62–47.

Villarreal, R. and Peterson, R.A. (2008) Hispanic ethnicity and media behavior. *Journal of Advertising Research*, **48** (2), 179–190.

internal marketing

Ashutosh Prasad

The term *internal marketing* (*IM*) appears to originate from Berry, Hensel, and Burke (1976). IM is recognized as an important component of a holistic marketing approach. It consists of marketing activities directed toward the employees of the firm as contrasted with the traditional, 'external' marketing.

If we conceptualize *employees as customers*, then IM can be seen as the shaping of their job products to satisfy their needs (Berry, 1981). Although the firm has quite a different relationship with its employees than with its customers, it is possible to find parallels to such external marketing activities as pricing (e.g., transfer pricing, compensation, bonuses), advertising (e.g., internal communications, training media), selling (e.g., face-to-face presentations), segmentation (e.g., team design, functional grouping), acquisition, and retention. This equivalence is appealing because it suggests that there is the potential to apply existing marketing techniques to the IM setting. However, extending the scope of marketing in this manner creates an overlap between the marketing and the personnel functions. To avoid interfunctional conflict, successful implementation of IM programs within the organization requires the support of upper-level management.

The proposal to treat employees as customers emerged from a services marketing (*see* SERVICES MARKETING STRATEGY) paradigm, with the implied or stated reasoning that satisfied employees would lead to satisfied customers. As George (1990, p. 63) notes, IM "... focuses on achieving effective internal exchanges between the organization and its employee groups as a prerequisite for successful exchanges with external markets."

Employee satisfaction is not an end in and of itself. Thus, a further conceptualization of IM might emphasize that the goal of IM is to increase customer satisfaction (*see* CUSTOMER SATISFACTION/DISSATISFACTION) by affecting employee satisfaction and other antecedent variables. This follows work, for example, by Grönroos (1985) that IM serves to "... motivate [employees] to customer-consciousness, market orientation and sales-mindedness by a marketing-like internal approach ..." (*see* MARKET ORIENTATION). Consistent with this, a textbook states that IM is simply "... training and motivating employees to serve customers well" (Kotler and Keller, 2009, p. 356), adding that, "... marketing functions must be coordinated from the customer's point of view ... [and] other departments must embrace marketing; they must also 'think customer.' Marketing is not so much a department as a company orientation" (Kotler and Keller, 2009, p. 24).

Several researchers have noted that IM can be a vehicle for implementing organizational change (Piercy, 1995; Rafiq and Ahmed, 2000). Rafiq and Ahmed (2000) note that in this role, IM is a planned effort using a marketing-like approach to overcome organizational resistance to change. Thus, a broad scope would be to treat IM as an organizational philosophy through which the firm becomes customer oriented. At the same time, it must be cautioned that while IM, through its employee focus, opens the door to using marketing theories to analyze a range of organizational issues, the fuzziness of its domain likely hampers its widespread adoption.

A large literature exists on IM theory and its use in corporations (Ahmed and Rafiq, 2002; Varey and Lewis, 2000). Prasad and Steffes (2002) discuss a scheme by Continental Airlines that sought to empower and reward employees for exceeding the industry average for on-time flight performance – a simple goal that would increase customer satisfaction and profits. The scheme was promoted by top management, and their active involvement and the obvious expense on IM may have convinced employees, who were initially distrustful of management, to "buy into" the scheme and through their collective involvement make it successful. In the absence of IM, failure could have become a self-fulfilling prophesy. Mitchell (2002) provides an example of what can happen when internal marketing is not used: "The failure of the merger of Deutsche Bank and Dresdner Bank in 2000 can be attributed in part to the failure of management to persuade Deutsche investment bankers of the vision for how the newly formed company would compete. Many key employees left and the threat of a mass walkout forced Deutsche to

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

abandon the deal after considerable damage to the share price of both companies."

Bibliography

Ahmed, P.K. and Rafiq, M. (2002) *Internal Marketing: Tools and Concepts for Customer-Focused Management.*, Butterworth-Heinemann, Oxford.

Berry, L.L., Hensel, J.S., and Burke, M.C. (1976) Improving retail capability for effective consumerism response. *Journal of Retailing*, **52**, 3–14.

Berry, L.L. (1981) The employee as customer. *Journal of Retail Banking*, **3**, 33–40.

George, W.R. (1990) Internal marketing and organizational behavior: a partnership in developing customer conscious employees at every level. *Journal of Business Research*, **20**, 63–70.

Grönroos, C. (1985) *Internal Marketing – Theory and Practice*. AMA Services Conference Proceedings, pp. 41–47.

Kotler, P. and Keller, K.L. (2009) *Marketing Management, 13E*, Pearson Prentice Hall, Upper Saddle River.

Mitchell, C. (2002) Selling the brand inside. *Harvard Business Review*, **80** (1), 99–104.

Piercy, N.F. (1995) Customer satisfaction and the internal market. *Journal of Marketing Practice*, **1** (1), 22–44.

Prasad, A. and Steffes, E. (2002) Internal marketing at continental airlines: convincing employees that management knows best. *Marketing Letters*, **13** (2), 75–91.

Rafiq, M. and Ahmed, P.K. (2000) Advances in the internal marketing concept: definition, synthesis and extension. *Journal of Services Marketing*, **16**, 449–462.

Varey, R.J. and Lewis, B.R. (2000) *Internal Marketing: Directions for Management*, Routledge, New York.

later mover (nonpioneer) advantage

Peter N. Golder

Later mover (or nonpioneer) advantage results from the challenges faced by first movers and the associated opportunities created for later movers. First movers face five key disadvantages relative to later entrants. In general, the primary challenges in converting a first-mover opportunity into long-term leadership are the high initial degree of market and technological uncertainty plus the dynamic competitiveness of all markets over time (*see* GO-TO-MARKET STRATEGY).

- *Free-rider effects*: First movers incur high costs to develop new products and the markets for those products. As a result, later entrants can sometimes acquire new technology and other inputs at a lower cost than the first mover. Diffusion of information across firms is one mechanism driving this process (*see* INNOVATION DIFFUSION). Similarly, later entrants may be able to hire more productive labor because workers already have experience with the new product.
- *Technological discontinuities*: New technologies create opportunities for later entrants to leapfrog first movers and possibly even make first movers' products obsolete. New technologies can lower costs and generate superior product attributes (*see* RADICAL INNOVATION). Adopting these newer technologies will be easier for later entrants when first movers have already invested heavily in the previous technology.
- *First mover misses consumers' ideal point*: In a new market, it is difficult to determine consumer preferences for the ideal combination of product attributes. If the first mover positions itself in a less desirable location, it may be too slow to reposition to the ideal point or the cost of repositioning may be too high. Meanwhile, a later entrant can learn from the first mover's mistake and position its product at the consumers' ideal point (*see* CUSTOMER SATISFACTION).
- *Shifting consumer tastes*: Successful first movers are able to identify and satisfy the needs of at least one market segment (*see* MARKET SEGMENTATION AND TARGETING). Their products become associated with their successful positioning. Yet, when consumer tastes shift, a first mover's past success becomes detrimental (*see* CONSUMER INNOVATIVENESS). A later entrant is better able to position its product at the new ideal point of consumers.
- *Incumbent inertia*: First movers may be unwilling to invest in maintaining their initial leadership. In some cases, this reluctance will occur for organizational reasons such as overcommitment to past means of success and unwillingness to cannibalize the current product line (*see* ORGANIZING FOR INNOVATION; RESEARCH & DEVELOPMENT). In other cases, first movers may have better opportunities in other markets and purposely implement a strategy of harvesting their market share.

Rewards for first movers are much lower than reported in studies initially supporting a first-mover advantage (*see* FIRST-MOVER (PIONEER) ADVANTAGE for more details). A comprehensive study of both surviving and failed first movers (identified through archival records rather than retrospective recollections) finds a survival rate for first movers of 36%, an average market share of 6% (*see* MARKET SHARE), and a market leadership rate of 9%. Nearly all first movers and those investing in such ventures expect much higher returns. The same study also finds that many so-called first movers are actually later entrants. Notable examples include Pampers, Gillette, Coca-Cola, and Amazon.com.

In order to overcome these disadvantages, first movers must seek to capitalize on all of the potential first-mover advantages. Often this effort requires firms to experiment with a variety of market and technology strategies until the optimal business model is identified and successfully implemented. Overall, both first movers and later entrants should not expect that entry timing alone will be a major factor in either firm's long-term success. Instead, all firms should focus on persistently developing their new products and new markets, and then relentlessly innovating and pursuing associated

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

market opportunities to capitalize on their past successes.

Bibliography

Golder, P.N. and Tellis, G.J. (1993) Pioneer advantage: marketing logic or marketing legend? *Journal of Marketing Research*, **30**, 158–170.

Shankar, V., Carpenter, G.S., and Krishnamurthi, L. (1998) Late mover advantage: how innovative late entrants outsell pioneers. *Journal of Marketing Research*, **35**, 54–70.

Tellis, G.J. and Golder, P.N. (2002) *Will and Vision: How Latecomers Grow to Dominate Markets*, McGraw-Hill, New York.

market definition

Glenn Voss

A product market is an exchange system defined by sellers, offerings, and customers. (*see* MARKET SEGMENTATION AND TARGETING; COMPETITOR ANALYSIS; SUPPLY CHAIN MANAGEMENT STRATEGY) In its simplest form, a product market is a single seller providing a single offering to a single customer. The complexity of the system increases dramatically as the number and diversity of competitors increase, as the number and diversity of offerings increase, and as the number and diversity of customers increase. At its most complex, a product market is synonymous with industry, with multiple competitors offering diverse portfolios of offerings that satisfy the diverse needs of a multi-segment market. This suggests that industry-level analytic frameworks such as Porter's 5 Forces (Porter, M.E. 1980) and product or industry life cycle (*see* STAGES OF THE PRODUCT LIFE CYCLE) can generate product market insights.

Meaningful product market analysis requires careful delineation of the target market (*see* MARKET SEGMENTATION AND TARGETING) and competitive offerings (*see* COMPETITOR ANALYSIS). Target market definition identifies the package of benefits customers are willing to purchase and includes an assessment of relative preferences for variations in benefit dimensions. Relevant competitive offerings include all technologies or solutions used to deliver the package of benefits customers seek. Direct competitors that use similar technologies to deliver comparable offerings constitute a single product market; indirect competitors that use different technologies to deliver substitute offerings may be considered as a distinct product market. This process may yield multiple segments within the product market or segments seeking sufficiently divergent benefits that warrant creation of separate product markets.

It is often useful to define the product market differently for different levels of strategic decision-making. Consider, for example, the worldwide market for cleaning products and one of the largest competitors, Procter & Gamble. At the level of corporate strategy, defining the product market as all cleaning products for all markets can generate synergies, especially from a supply and production perspective. This macrolevel product market definition also facilitates identifying global demand trends as well as all potential competitors and substitute technologies.

However, to make more effective business and marketing decisions, P&G splits the product market in several ways and organizes and manages its offerings accordingly. First, there is a distinction between business markets and consumer markets. Some of the offerings for the two markets are the same (e.g., Tide and Cheer) and some are unique (e.g., P&G Pro Line Floor Finish Stripper). The purchase and use behavior of the business and consumer markets is sufficiently different to warrant distinct business and consumer divisions responsible for packaging, promotion, sales, pricing (*see* PRICING STRATEGY), channel, (*see* SUPPLY CHAIN MANAGEMENT STRATEGY) and CUSTOMER RELATIONSHIP MANAGEMENT efforts. Within each division, there are additional product market subdivisions. For example, consumer product markets are defined for cleaning dishes and cleaning clothes, among others. The dishwashing product market can be subdivided again into automatic dishwashing needs and offerings, and hand dishwashing needs and offerings. Again, market needs are sufficiently different to warrant distinct consideration of positioning (*see* POSITIONING ANALYSIS AND STRATEGIES), packaging, promotion, and pricing (*see* PRICING STRATEGY) efforts.

Geographic markets and types of channels (*see* SUPPLY CHAIN MANAGEMENT STRATEGY) used to reach end customers also can be useful in defining product markets. For example, soft drink companies like Coca-Cola and Pepsi frequently develop different offerings that satisfy the beverage preferences of different global markets. Coca-Cola and Pepsi create different packaging and delivery offerings for the quick-service restaurant channel (fountain drinks), the supermarket channel (larger packages), and the convenience store and vending channel (single-serving offerings). In each case, customer needs (*see* CUSTOMER ANALYSIS) and competitive offerings (*see* COMPETITOR

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

ANALYSIS) may be sufficiently varied to warrant distinct product market consideration.

See also *competitor analysis; customer analysis; market segmentation and targeting; positioning analysis and strategies; pricing strategy; stages of the product life cycle; supply chain management strategy*

Bibliography

Porter, M.E., (1980), *Competitive Strategy*, Free Press, New York.

market orientation

Stanley F. Slater, Jakki J. Mohr, and Sanjit Sengupta

In this article on market orientation we highlight the role of market-based information – about customers, competitors, and other important stakeholders and trends – in strategic decision making. Simply put, information about the market becomes the beacon to guide the company through the messy, oftentimes political, process of strategy making. With a shared understanding of the value and use of market intelligence, a company can create a powerful knowledge-based competency that allows it to gain advantage in the marketplace.

What Does it Mean to be Market Oriented?

Businesses that value close customer relationships and rely on market information to guide strategic decision making are commonly described as market oriented. Market-oriented businesses generate intelligence about customers' current *and* future needs, and about competitors' capabilities and strategies; share that intelligence throughout the organization; and take coordinated action to create superior customer value (e.g., Kohli and Jaworski, 1990; Narver and Slater, 1990).

The effect of market orientation on company performance. A significant body of research has demonstrated a positive relationship between the degree of a firm's market orientation and its performance (*see* Kirca, Jayachandran, and Bearden, 2005 for an extensive review and synthesis). As shown in Figure 1, by utilizing market intelligence to continuously improve both product quality and service quality, and develop innovative products that meet evolving customer needs, a market orientation delivers superior sales growth and profitability.

Some research has found that the positive relationship between market orientation and firm performance holds true regardless of the environmental conditions in which a firm competes (such as high or low levels of market turbulence, technological turbulence, competitive intensity, and market growth rates) (Slater and Narver, 1994); however, other studies report that the market orientation/firm performance relationship is *stronger* in highly dynamic markets (Homburg and Pflesser, 2000). Firms must excel not only at generating new innovations that deliver value to customers but also at commercializing these innovations (*see* LAUNCH STRATEGIES). A strong market orientation without commensurate development of a strong innovation/technological capability (*see* CORE COMPETENCIES) can have a negative effect on new product and market performance (Baker and Sinkula, 2005).

Activities of the market-oriented business.

Intelligence generation. First, market-oriented firms generate a wide array of intelligence about influential market forces. Market intelligence includes useful information about trends and participants in the market including

- current and future customer needs (*see* CUSTOMER ANALYSIS);
- competitors' capabilities and strategies (*see* COMPETITOR ANALYSIS); and
- other important influencers, such as suppliers, as well as important market trends such as socioeconomic and technological change.

Customer intelligence may be generated through traditional market research techniques such as focus groups, customer surveys, customer advisory boards, and feedback from the salesforce (*see* FOCUS GROUPS AND DEPTH INTERVIEWS). However, the intelligence generated through these processes provides insight only into needs that customers can easily articulate. Responding to customers' expressed needs is usually inadequate to create superior, sustainable customer value since customers for many products may find it very difficult to clearly articulate their needs; they may be unaware of the capabilities that new technologies offer or lack insight into how technology can address their needs. And, even if customers are able to articulate their needs, what they articulate to one seller can and will be articulated to competitive sellers as well.

Firms that primarily act on current intelligence are said to practice a *responsive market*

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

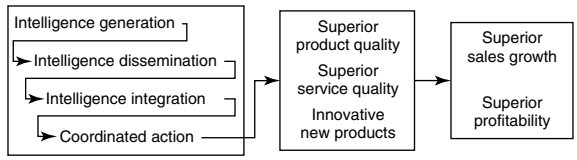


Figure 1 Market Orientation, Positional Advantage, and Performance.

orientation. Although current intelligence provides important information, some argue that listening to current customers too carefully can inhibit innovation, constraining it to ideas that customers can envision and articulate – which may lead to safe, but bland, offerings. Firms with a responsive market orientation may have a tendency to focus very specifically on solving existing customers' needs with a current technology. Firms that focus too narrowly on their established customers may be constrained in the strategies and technologies they choose to pursue. Such a myopic focus obscures the possibility that customer needs may change over time and may be solved in radically different ways, allowing new, disruptive innovations to creep up like a stealth attack.

For example, a firm's largest and most profitable customers may be the last to embrace a radical innovation and the company may believe that the people who embrace a new technology first, say a simple or inexpensive new technology, may be an unattractive market segment – and as a result, the company pays it little attention. Take the 5.25-inch disk drive, a technology introduced in the early 1980s that was embraced by the emerging desktop personal computer marketplace – yet was inconsistent with mainframe and minicomputer customer demands (Christensen and Bower, 1995). Established disk drive firms did not fail because they were unable to develop innovative technologies. Rather, because established customers were uninterested in new technologies that didn not address their immediate needs, industry leaders did not allocate resources to developing the new technologies. This decision allowed new entrants to gain leadership in the new market.

Businesses that gather anticipatory intelligence have a *proactive market orientation*. Anticipatory customer intelligence is concerned with customers' latent and future needs that enable the firm to proactively pursue market opportunities that are not evident to competitors. Latent needs are real needs that are not yet in the customers' awareness. If these needs are not satisfied by a provider, there is no customer demand or response. Customers are not necessarily dissatisfied, because the need is unknown to them. However, if a company understands such a need and fulfills it, the customer is "wowed," and rapidly delighted. Offering products and services that address these latent needs in a compelling fashion delights and excites customers and inspires loyalty. A structured customer visit program, ethnographic/observational research, and working with lead users, customers whose needs are more advanced relative to the rest of the market, may provide insight into customers' latent and future needs (*see* OBSERVATION METHODS; ETHNOGRAPIC RESEARCH; PERSONAL OBSERVATION; LEAD USERS). While a responsive market orientation may generate incremental innovations, a proactive market orientation is more likely to lead to RADICAL INNOVATION; (Atuahene-Gima, Slater, and Olson, 2005). Therefore, firms must be ambidextrous, combining both responsive and proactive market orientations.

An extreme type of proactive market orientation is known as *market driving* in which a firm actively seeks to (i) redefine the structure of the market (*see* MARKET/INDUSTRY STRUCTURE) and/or (ii) introduce an innovative value proposition (*see* A FRAMEWORK FOR CREATING VALUE PROPOSITIONS; VALUE PROPOSITION) that enables the firm to reduce, or even avoid, competition (Jaworski, Kohli, and Sahay, 2000). Market-driving activities may be focused on many different stakeholders including customers and competitors, as well as vendors, potential partners and allies, and regulators. A market-driving strategy can be quite risky since managers are attempting to change

the structure of a market and/or the behavior of players in the market by introducing a discontinuous leap in the customer value system supported by a unique business system. For example, Apple anticipated the emergence of the PDA (personal digital assistant) market when it introduced the Newton in 1993. The Newton Message Pad featured a variety of personal-organization applications, such as an address book, a calendar, notes, along with communications capabilities such as faxing and email. It featured a pen-based interface which used a word-based, trainable handwriting recognition engine. Unfortunately, this engine was notoriously difficult to use which, along with competition from the PalmPilot, led to its cancellation in 1998.

Competitive intelligence can be gleaned through a wide variety of online and off-line resources. The Society of Competitive Intelligence Professionals (http://www.scip.org) and the Competitive Intelligence Resource Index (www.bidigital.com/ci/) are two useful starting places. Competitive intelligence can include information on competitors' customer lists, product and pricing information, new-product plans and R&D efforts, job postings that provide insights into new business arenas, information about partnerships, alliances, distributors, and managers, to name a few.

While both customer- and competitor-oriented intelligence generating activities are important, a customer orientation, particularly a proactive orientation, is more strongly associated with a firm's ability to develop new knowledge and skills (*see* MARKET-BASED ASSETS), which, in turn, is more likely to produce breakthrough product and process innovations (*see* INNOVATION TYPOLOGIES). Conversely, competitor-oriented businesses are more likely to invest resources into improving existing products, and hence are more likely to produce incremental innovations.

Regardless of the specific tools used to gather market-based information, it is imperative that the firm allocates resources to the information-gathering process. However, this is much easier said than done. Many firms do not give adequate resources to initiatives that seemingly are not directly tied to underlying product development, commercialization, and relationship building. Indeed, one $2 billion US-based high-tech company with which we are familiar was unwilling to invest any meaningful funds to purchase secondary market data, analyst reports, or to undertake its own primary market research. Yet this company wanted to make a "cultural transformation" to being market driven. Clearly, a company that says it wants to become more customer focused must put its money where its mouth is.

Intelligence dissemination. As important as intelligence generation is, the information generated will be of limited value until it is shared across the organization and potentially combined with other information. A firm's knowledge-based competitive advantage – one that increasingly resides in its know-how – is only as strong as its ability to share and use knowledge within and across the organization's boundaries (*see* CROSS-FUNCTIONAL TEAM). Indeed, some say that effective knowledge management requires a boundaryless organization, which takes good ideas from disparate functions and from outside organizations and uses them in many areas. Indeed, as stated by Lew Platt, former CEO of Hewlett-Packard, "*If only we knew what we know*" (cited in Brown and Duguid, 2000). These words reflect a tough truth about most organizations: the knowledge and the know-how of their workforce is too often underused – isolated in departments and functional units.

Market-oriented firms actively encourage sharing of information across people, departments, and divisions. Information may be shared formally in meetings, conferences, newsletters, and databases, or informally through "hall talk" and informal mentoring. People in the organization must be able to ask questions and augment or modify the information to provide new insights to the sender.

However, freely sharing and using information is easier said than done. An organization is a coalition of individuals or groups, each of which has its own goals that may conflict with organizational goals. One way that individuals or departments promote their own self-interests is by creating and protecting their "proprietary" knowledge base. They believe that hoarding or withholding information protects their turf and generates power and status. However, this

dysfunctional belief does not allow the organization to fully leverage its knowledge. Effective dissemination increases the value of information, allowing each piece of information to be seen in its broader context by all organizational players who might be affected by it, use it, or possess complementary information about it.

Thus, the challenge is to create an organizational environment in which group success does not come at the expense of individual success. One way to ensure this is to cultivate a team orientation as the foundation for effective information sharing, a topic that we address later in this article.

Intelligence integration. Third, market-oriented companies must integrate intelligence to create knowledge assets, and they must achieve a shared belief system: a shared interpretation of the information and its implications for the business. Shared interpretation goes far beyond simply sharing the information. Consider this variation on an old joke:

> *Marketing:* The glass is half full.
>
> *R&D:* The glass is half empty.
>
> *Operations:* The glass is twice as large as it needs to be.

Each of these perspectives carries substantially different views of what a particular piece of information means, and, by extension, implications for what the firm should do. Integrating intelligence in order to achieve a shared understanding is vital.

Indeed, firms in dynamic and complex markets must strive for consensus internally in order to successfully develop and execute strategy. However, prior to achieving consensus, companies in volatile industries may benefit from a relatively high level of disagreement among managers in interpreting the information they have gathered. Such disagreement allows a closer inspection of the validity of different assumptions and alternatives, as well as an assessment of the relative importance of the company's objectives and competitive methods. In his 1999 book, *Only the Paranoid Survive*, Intel's chairman Andy Grove described his view of spirited debate to his middle managers: "Your criterion for involvement should be that you are heard and understood .... All sides cannot prevail in the debate, but all opinions have value in shaping the right answers." A company must reach a shared interpretation of the information it has gathered, but it should not do so prematurely. It should actively facilitate debate, discussion, disagreement, and dialogue in order to fully tap the value of its knowledge.

Coordinated action for implementing decisions. Finally, the market-oriented firm implements decisions through coordinated action. Commitment to execution is necessary for successful implementation of a market orientation. An organization can generate and disseminate intelligence; however, unless it acts on that intelligence, nothing will be accomplished. Mark Hurd, CEO of Hewlett-Packard, quoting Einstein, observed (Murray, 2007),"Vision without execution is hallucination." Acting on market intelligence involves making decisions including the selection of target markets; the development of products/services that address customers' current and anticipated needs; and production, distribution, and promotion of the products to engender both customer satisfaction and customer loyalty (Kohli and Jaworski, 1990).

Moreover, the idea of coordinated action means that all functions in a market-oriented company – not just marketing – participate in responding to market needs. To drive new products from concept to launch more rapidly and with fewer mistakes, "all functional interfaces (contacts) must jointly share in discussions and information exchange" (Gupta, Raj, and Wilemon, 1986). Coordinated action requires a greater emphasis on multifunctional activities – activities that are the joint responsibility of multiple functions in the business. When decisions are made interfunctionally and interdivisionally, a more accurate representation of the information and a closer connection to the market issues will occur. Moreover, interfunctional decision making implies that the people who will be involved in implementing the decisions are the ones actually involved in making the decisions – the idea being that involvement in making the decision makes one more committed to implementing that

decision. Indeed, two firms may have the same information, but what allows one firm to leverage that knowledge more successfully than another is not the knowledge itself, but rather, how that knowledge is coordinated and integrated among different functions and departments.

However, interfunctional rivalry can impede internal knowledge transfer. Functions must compete for scarce organizational resources and are thus often reluctant to share information. This sets up a situation sometimes referred to as *coopetition*, or the simultaneous need for cooperation and competition by various departments within a company. It is true that functional areas compete with one another for resources, but they also cooperate to work toward the firm's common interests. A study of these complicated intraorganizational dynamics found that fostering cooperation but squelching internal competition can actually limit a firm's performance potential (Luo, Slotegraaf, and Pan, 2006). Constructive conflict can promote learning and knowledge sharing, and improve performance. Importantly, market learning is the mechanism by which the simultaneous effects of interdepartmental cooperation and competition are transferred to higher performance. Hence, problems in a firm's market orientation may be due not only to low cross-functional cooperation but also to an inability to assimilate and deploy market knowledge.

In sum, a firm's level of market orientation is determined by its ability to generate and disseminate market intelligence, to come to shared understanding of what that information means, and to make and execute decisions based on the information with cross-functional representation, including multiple departments that are affected by the decisions. How does a company know how market oriented it is? The following section provides guidelines to assess the degree of market orientation of a firm.

Assessing a firm's degree of market orientation. Table 1 contains a set of statements that describe market-oriented business practices. A manager should rate his/her business on each item using the following scale:

Strongly Disagree (−3)	Disagree Moderately (−2)	Disagree Slightly (−1)
Agree Slightly (1)	Agree Moderately (2)	Strongly Agree (3)

Then sum the scores for the four items in each of the categories. Negative scores indicate much room for improvement, low positive scores indicate the need for selective or incremental improvement, while high scores indicate a strong capability. High scores are relatively rare.

Becoming market oriented: facilitating conditions. The logic for and evidence of the value of a market orientation are clear and compelling. Why, then, are more firms not able to become market oriented? Achieving a market orientation requires a cultural change in the organization. It requires shared values and beliefs about the need to gather, share, and compile market-based information; it requires a common desire to actively discuss the meaning of that information, and to ensure that the company's decisions are soundly grounded in the relevant information. It also requires a resource commitment to the processes involved in a market orientation: data collection expenditures and staffing, time to share and discuss the meaning of the information, and so forth.

The fact is, many firms do not achieve a market orientation because they do not have the proper facilitating conditions in place. These conditions include, at a minimum (i) prioritization of a firm's information gathering focus to match its strategy, (ii) top management advocacy, (iii) a flexible, decentralized structure, (iv), a team orientation, and (v) a market-based reward system.

Prioritizing information gathering needs. Resource constraints coupled with highly complex markets make it impossible to comprehensively scan complex environments. To make sense of complex environments, managers must focus their scanning efforts on the market forces that are most salient to their strategies (*see* MARKETING STRATEGY).

Table 1 Market-oriented business practices.

Responsive Customer Intelligence Generation:
- We continuously work to better understand our customers' needs.
- We pay close attention to after-sales service.
- We measure customer satisfaction systematically and frequently.
- We want customers to think of us as allies.

Responsive Competitor Intelligence Generation:
- Employees throughout the organization share information concerning competitors' activities.
- Top management regularly discusses competitor's strengths and weaknesses.
- We track the performance of key competitors.
- We evaluate the strengths and weaknesses of key competitors.

Proactive Customer Intelligence Generation:
- We continuously try to discover additional needs of our customers of which they are unaware.
- We incorporate solutions to unarticulated customer needs in our new products and services.
- We brainstorm about how customers' needs will evolve.
- We work with lead users, customers who face needs that eventually will be in the market – but face them months or years before the majority of the market.

Proactive Competitor Intelligence Generation:
- We try to anticipate the future moves of our competitors.
- We monitor firms competing in related product/markets.
- We monitor firms using related technologies.
- We monitor firms already targeting our prime market segment but with unrelated products.

Intelligence Dissemination:
- We have interdepartmental meetings to discuss market trends and developments.
- Marketing personnel spend time discussing customers' needs with other functional departments.
- We share information about major market developments.
- Data on customer satisfaction are shared at all levels in the organization.
- When one function acquires important information about customers or competitors, it shares that information with other functions.

Intelligence Integration:
- We have cross-functional meetings for the purpose of intelligence integration.
- We reach organizational consensus regarding the holistic meaning of related pieces of information before taking action.
- We utilize cross-functional teams or task forces for important initiatives to ensure that all points of view are considered before decisions are made.
- Collaboration is valued in this business.

Coordinated Action:
- We are quick to take advantage of market opportunities.
- The activities of different functions in this business are well coordinated.
- We make sure that all critical functions understand our objectives and strategy before we take action.
- There is a high level of cooperation and coordination among functional units in setting the goals and priorities for the organization to ensure effective response to market conditions.

Without the ability to simplify, structure, and focus their intelligence generation efforts, managers would suffer from "paralysis by analysis." One approach to prioritizing their market-scanning efforts is to identify the critical issues faced by each of four different strategy types. Depending upon whether a company's strategy is a Prospector (product leader), Analyzer (fast follower), Low-Cost Defender (operationally excellent) or a Differentiated

Defender (customer intimate), it will collect and use market-based information differently (Olson, Slater, and Hult, 2005).

Prospectors continuously seek to locate and exploit new product and market opportunities. Owing to the importance of product-market innovation, the most successful Prospectors emphasize proactive customer intelligence generation and generation of intelligence about the development of innovations. Such intelligence can lead either to product performance improvements for current customers or to new products that penetrate new markets.

Defenders attempt to seal off a portion of the total market to create a stable set of products and customers. *Differentiated Defenders* accomplish this by providing superior service and/or product quality. *Low-Cost Defenders* focus on producing goods or services as efficiently as possible so that they can provide their products or services at the best prices. Because the value proposition for the Low-Cost Defender entails producing their products/services as efficiently as possible and offering them at a relatively low price, the most successful Low-Cost Defenders focus their intelligence generation efforts on process rather than on product innovation and on competitors – competitor intelligence provides a benchmark against which prices, costs, and performance can be compared. Differentiated Defenders focus on retaining customers through attention to superior service, product quality, or image. Consequently, the most successful Differentiated Defenders emphasize the generation of customer intelligence and the establishment of strong customer relationships.

Analyzers occupy an intermediate position by cautiously following Prospectors into new product-market domains while simultaneously protecting a stable set of products and customers. In order to identify opportunities for potential product improvements or for entering new market segments, Analyzers must closely monitor customer reactions to Prospectors' offerings as well as competitors' activities, successes, and failures. In other words, while developing customer intelligence is essential for Analyzers, monitoring competitors' actions may be as important to their success.

Top management commitment. If a firm's top managers are not unequivocally and visibly committed to its customers and the collection of market-based information, then the firm will not bring its resources to bear on developing solutions to meet its customers' needs (Jaworski and Kohli, 1993). As Michael Dell, CEO of Dell Computers, once said (Mears, 2003):

> "We have a relentless focus on our customers. There are no superfluous activities here. Once we learn directly from our customers what they need, we work closely with partners and vendors to build and ship relevant technologies at a low cost. Our employees feel a sense of ownership when they work directly with customers."

Decentralized organizational structure. Market-oriented behavior thrives in an organization that is decentralized with fluid job responsibilities and extensive lateral communication processes (Jaworski and Kohli, 1993). Members of these organizations recognize their interdependence and are willing to cooperate and share market intelligence to sustain the effectiveness of the organization. Effective information sharing in a market-oriented organization demands that bureaucratic constraints on behavior and information flow be dismantled. Market uncertainty requires high frequency and informality in communication patterns among organizational units for effective intelligence dissemination.

Team orientation. Team orientation is essential for effective intelligence sharing. In a team-oriented organization, value is placed on working cooperatively toward common goals for which all employees feel mutually accountable. The organization relies on teamwork to get work done. Teamwork has several contributing processes including (i) conflict management, (ii) motivation and confidence building with regard to the team accomplishing its goals and objectives, and (iii) "affect management" which represents those activities that foster emotional balance, togetherness, and effective coping with stressful demands and frustration. Team effectiveness is enhanced when teams monitor progress toward goals, track resources necessary to achieve the team's goals, team members provide direct or indirect feedback to other team

members, and coordinate the activities of team members.

Market-based compensation system. By rewarding employees for generating and sharing market intelligence and for achieving high levels of customer satisfaction and customer loyalty, a market-based reward system is the single organizational factor that has the greatest impact on a market orientation (Jaworski and Kohli, 1993). Market-oriented firms place less emphasis on short-term sales and profit goals than their more financially or internally focused competitors.

At its heart, creating a market-oriented organizational culture within a firm requires a major transformation, characterized by four stages: initiation, reconstitution, institutionalization, and maintenance (Gebhardt, Carpenter, and Sherry, 2006). Essentially, members of the company recognize a threat. To address the threat, a group of empowered managers creates a coalition to mobilize the larger organization; they create a process that reconnects the company's personnel with its customers. Through this process, personnel build common experiences and perspectives and build a consensus for more formal organizational changes to solidify the cultural shift. These formal changes then sustain the new orientation.

Summary

At the most basic level, market-oriented firms excel at generating, sharing, and using information about the market (customers, competitors, collaborators, technology, trends, etc.) to make coordinated decisions that lead to the creation of innovative new products, and superior product and service quality, which lead to superior performance. However, some firms are more proactive than others both in the type of intelligence they generate and in the action they take. The more proactive businesses seek intelligence not only about current customer needs and competitor threats but also about latent and future customer needs and competitive threats. Proactive businesses are more likely to develop market-driving strategies in which a firm actively seeks to redefine the structure of the market, and/or introduce an innovative value proposition that enables the firm to reduce, or even avoid, competition. A market-driving strategy, while riskier than a responsive strategy, is more likely to lead to a breakthrough position. To successfully transition to being a market-oriented organization, there must be strong leadership from top management, decision making that takes place as close to the customer as possible, and a reward system that recognizes the importance of information sharing and coordinated action.

Bibliography

Atuahene-Gima, K., Slater, S., and Olson, E. (2005) The contingent value of responsive and proactive market orientations for product Innovation. *Journal of Product Innovation Management*, **22** (6), 464–482.

Baker, W. and Sinkula, J. (2005) Market orientation and the new product paradox. *Journal of Product Innovation Management*, **22** (6), 483–502.

Christensen, C. and Bower, J. (1995) Customer power, strategic investment, and the failure of leading firms. *Strategic Management Journal*, **17**, 197–218.

Brown, J.S. and Duguid, P. (2000) *The Social Life of Information*, Harvard business school press, Boston.

Gebhardt, G., Carpenter, G., and Sherry, J. (2006) Creating a market orientation: a longitudinal, multi-firm, grounded analysis of cultural transformation. *Journal of Marketing*, **70** (4), 37–55.

Gupta, A., Raj, S., and Wilemon, D. (1986) A model for studying the R&D – marketing interface in the product Innovation process. *Journal of Marketing*, **50** (2), 7–17.

Homburg, C. and Pflesser, C. (2000) A multiple-layer model of market-oriented organizational culture: measurement issues and performance outcomes. *Journal of Marketing Research*, **37** (4), 449–462.

Jaworski, B., Kohli, A., and Sahay, A. (2000) Market-driven versus driving markets. *Journal of the Academy of Marketing Science*, **28** (1), 45–54.

Jaworski, B. and Kohli, A. (1993) Market orientation: antecedents and consequences. *Journal of Marketing*, **57** (3), 53–70.

Kirca, A., Jayachandran, S., and Bearden, W. (2005) Market orientation: a meta-analytic review and assessment of its antecedents and impact on performance. *Journal of Marketing*, **69** (3), 24–41.

Kohli, A. and Jaworski, B. (1990) Market orientation: the construct, research propositions, and managerial implications. *Journal of Marketing*, **54** (2), 1–18.

Luo, X., Slotegraaf, R., and Pan, X. (2006) Cross-functional 'Coopetition:' the simultaneous role of cooperation and competition within firms. *Journal of Marketing*, **70**, 67–80.

market evolution

Glenn Voss

A product market is a dynamic exchange system that evolves over time. The number of competitors (*see* COMPETITOR ANALYSIS), the diversity and number of product offerings, and the size and needs of the customer market (*see* CUSTOMER ANALYSIS) are in continuous flux. The industry or product life cycle concept (*see* STAGES OF THE PRODUCT LIFE CYCLE) is one of the best-known frameworks for considering product market evolution. However, the life cycle concept proposes a deterministic progression from introduction to growth to maturity and decline, with the number of firms initially increasing, then decreasing, stabilizing, and ultimately declining. This contrasts with the observation that some product markets seem to remain forever fragmented and with the idea that a product market is a dynamic exchange system.

As dynamic systems, product markets exist along a continuum ranging from relatively stable, incremental evolution to frequent, dynamic, and radical change. Market dynamism is driven by interactions between sellers who experiment with products and prices to increase market share and customers who have heterogeneous and dynamic preferences. These interactions create feedback loops, so that changes in product offerings can lead to altered buyer preferences, which in turn can lead to additional seller innovation and heterogeneity.

Innovation and imitation play key roles in product market evolution. Successful innovation provides advantage to the innovating firm (*see* FIRST-MOVER (PIONEER) ADVANTAGE), with the duration of the advantage largely dependent on the ability of competing firms to imitate the innovation. Where innovation is easy to imitate, imitators tend to outperform innovators because their lower overall cost structure allows for greater margins or lower prices (*see* LATER MOVER (NONPIONEER) ADVANTAGE). However, industry concentration does not occur as long as profitable, imitative firms show restraint in increasing output, which is likely in the absence of economies of scale or scope. Thus, when imitation is easy and economies of scale and scope are modest, both innovation and imitation can be profitable and product markets can remain fragmented indefinitely (e.g., restaurants and hair stylists).

Product markets tend to consolidate and stabilize when economies of scale or scope are present. If innovation is easy to imitate, imitative firms build scale and scope advantages into dominant positions (e.g., Walmart). Where innovation is difficult to imitate, industries become concentrated as successful innovators (or lucky early imitators) leverage their advantage to build dominant market share (e.g., electronics). As concentration increases, the remaining firms maximize profits by maintaining high margins and limiting capacity increases. The result is a relatively stable oligopoly where dynamism and variety are lower when compared to more fragmented product markets. This is consistent with the maturity stage of the product life cycle (*see* STAGES OF THE PRODUCT LIFE CYCLE).

Oligopolists in stable markets use product offering proliferation to satisfy diverse customer preferences. However, when entry barriers are low and customers seek variety, innovators can enter stable markets with new offerings. For example, the top three brewers dominated the US beer market for decades. And while the top three continue to maintain dominant market share, thousands of craft brewers have entered the market over the last two decades, winning small slices of a large market. Thus, over the last 100 years, the US beer market has evolved from thousands of regional and independent brewers to a small number of dominant oligopolists competing amongst themselves and with thousands of independent craft brewers. This is because some beer drinkers value variety and are innovative enough to support the efforts of small, innovative craft brewers.

See also *competitor analysis; customer analysis; first-mover (pioneer) advantage; later mover (nonpioneer) advantage; stages of the product life cycle*

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

market segmentation and targeting

Shikhar Sarin

Introduction

The concepts of segmentation and targeting lie at the heart of marketing strategy (Best, 2004). In the following discussion, we define and present the rationale for segmentation and targeting, outline the criteria used for evaluating effective segmentation, discuss the commonly used bases for segmenting consumer and business markets, and conclude by describing various approaches for targeting segments.

Definitions and Rationale

Markets are heterogeneous. They consist of customers with varying needs and preferences. It is very difficult to design a single product or a MARKETING MIX that will satisfy all the customers in a market (Boone and Kurtz, 2001). Marketers therefore divide the total market into smaller, relatively homogeneous groups. This process of dividing heterogeneous markets of dissimilar customers into smaller, more homogeneous subgroups of customers having similar tastes, preferences, and/or needs is called *segmentation* (Boone and Kurtz, 2001; Zikmund and d'Amico, 2001). These resulting subgroups are referred to as *market segments*. Wind and Cardozo (1974) define a market segment as a group of current or potential customers with some common characteristic that is relevant in planning and predicting their response to a firm's marketing stimuli.

Owing to resource constraints, most firms are unable to pursue all the segments they identify in a given market. A firm is likely to focus its marketing efforts on a smaller number of meaningful marketing segments, and concentrate on satisfying these selected segments of the total market (Zikmund and d'Amico, 2001). Zikmund and d'Amico describe targeting as the process of allocating resources effectively by focusing marketing efforts on a selected part of the total market.

Hutt and Speh (2010) note that market segmentation and targeting increase a firm's chances of success by selecting a well-defined group of potentially profitable customers; focusing marketing resources on these customers; and developing a distinct value proposition (*see* A FRAMEWORK FOR CREATING VALUE PROPOSITIONS; VALUE PROPOSITION) that meets the needs of these customers (in terms of product/service and/or marketing mix) better than the competition (*see* COMPETITIVE ADVANTAGE).

Criteria for Effective Segmentation

Segmentation by itself does not ensure success. In general, four criteria are used to evaluate the effectiveness of segmentation, or the desirability of potential market segments (Hutt and Speh, 2010; Boone and Kurtz, 2001):

1. *Measurability* – the degree to which information on the segment characteristics exists, or can be obtained.
2. *Accessibility* – the degree to which marketers can effectively promote to and serve the market segments.
3. *Substantiality* – the degree to which the segments are large or profitable enough to merit consideration for their own marketing mix.
4. *Responsiveness* – the degree to which the segments respond differently to different elements of the marketing mix (i.e., pricing, product, etc.). Some marketers consider reliability of a segment's response to the marketing mix as an important part of responsiveness as well.

Segmentation attempts to isolate the characteristics that distinguish the buying behavior of a certain group of customers from other groups, or from the overall market (Boone and Kurtz, 2001). These characteristics form the bases for segmentation. Customers can either be individuals (i.e., consumer markets) or organizations (i.e., business markets), and the bases for segmentation used for each market differ significantly. In the following discussion, we cover some of the commonly used bases for segmenting consumer and business markets.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Bases for Segmenting Consumer Markets

Bases for segmenting consumer markets are broadly classified under four major categories: geographic, demographic, psychographic, and product-related segmentation (Boone and Kurtz, 2001). Geographic and demographic segmentation are more popular because of their relative ease of use. While information on psychographic and product-related segmentation is harder to come by, these bases of segmentation are considered to be more effective.

Geographic segmentation. Geographic segmentation is one of the most popular approaches to segmenting consumer markets. Marketers often use geography to divide markets when needs and preferences of customers differ based on variables like political or cultural boundaries, climatic regions, and so on. For example, in countries like the United Kingdom and India cars have right-hand drive, whereas they have left-hand drive in countries like the United States. It is not uncommon for global automobile manufacturers like Toyota or Honda to develop a left-hand drive and a right-hand drive model for each car they make to be able to cater to the needs of their consumers in different locations.

People also differ in their preferences across countries and cultures. In Argentina, most Coca Cola is consumed with food, whereas in most Asian countries Coca Cola is consumed independently as a refreshing beverage, and rarely served with meals (Zikmund and D'Amico, 2001). Smart marketers recognize these differences and fine-tune their product offerings and marketing mix accordingly. For example, European countries tend to have smaller living quarters and higher utility rates. Hence, washing machines developed for the European markets often have a smaller capacity but more efficient motors than those sold in the US market, where home sizes tend to be larger and the utility rates lower – resulting in a consumer preference for bigger appliances.

Commonly used sources for geographic information include the census data from various national governments, and Geographical Information Systems (*see* Boone and Kurtz, 2001).

Demographic segmentation. This kind of segmentation involves defining consumer groups based on demographic variables such as age, income, gender, education, occupation, household type, race, and family lifecycle (Boone and Kurtz, 2001). Some illustrative examples are discussed below.

Gender. It has long been assumed that consumer electronics are a male-dominated purchase, with a lot of the advertising for such goods emphasizing technical jargon (Boone and Kurtz, 2001). However, a recent survey by the Consumer Electronics Marketers Association (CEMA) revealed that women are responsible for over half of the consumer electronic purchases. Female consumers look at electronics to serve a practical purpose, to be a tool or a source of entertainment; they find the technical jargon in most consumer electronic advertisements unappealing. On the basis of such information, companies like Microsoft refined their advertising and promotion strategy for women, de-emphasizing the technical jargon and emphasizing the benefits delivered by their products (Boone and Kurtz, 2001).

Age. Online social networking (*see* SOCIAL NETWORKS) sites are very popular nowadays. Many of these sites segment their users based on age. Two of these sites, Facebook and MySpace, target very different kind of users. Facebook users generally tend to be in their 30s, 40s, and older, whereas MySpace tends to target younger users by concentrating their recruitment and promotional efforts on college- and school-age populations.

Family lifecycle. Most adults move through different stages in life such as being single, being married, having kids, raising kids, living together after kids have left the house, and so on. The family lifecycle represents the traditional stages through which most families pass (Zikmund and d'Amico, 2001). Each stage will have specific needs, preferences, and consumption patterns associated with it. For example, first-time parents tend to be heavy consumers of baby items such as cribs, strollers, and diapers. Consumption patterns of parents with kids at home are heavily influenced and driven by the kids (*see* FAMILY BUYING). On

the other hand, single individuals tend to be heavy users of personal care items and consumer electronics (Boone and Kurtz, 2001). Marketers often use stages of family lifecycle to segment their consumers.

Household type. Census data show that across many developed countries, the number of traditional households (i.e., married with kids) is decreasing, whereas the number of nontraditional households such as single parents, people of opposite sex in same living quarters, and double income couple with no kids (DINKS) is on the rise. Each type of household exhibits a specific profile. For example, DINKS have high levels of disposable income, and tend to be big buyers of gourmet foods, luxury items, and travel (Boone and Kurtz, 2001), making them a very attractive segment for such products.

Because data on many of the demographic variables are relatively easy to obtain, demographic segmentation remains one of the most popular methods of segmenting markets. Zikmund and d'Amico (2001) note that some commercially available systems like PRIZM use geographic variables like mail codes to cluster demographically similar individuals to create geodemographic segmentation. While the census information collected varies greatly from country to country, the following sources can offer population statistics for many countries: US Bureau of the Census – International Program Center, United Nations – National Statistics Offices, and Euromonitor – a private market research firm specializing in demographics of European residents (Boone and Kurtz, 2001).

Psychographic segmentation. Psychographic segmentation divides a population into groups that have similar psychological characteristics, values, and lifestyles (Boone and Kurtz, 2001). Lifestyle refers to an individual's life goals – how the person spends his/her time and money (Zikmund and d'Amico, 2001). Consumer's lifestyles are composites of their psychological profiles, needs, motives, perceptions, and attitudes (Boone and Kurtz, 2001).

Quantitative measures of such psychological characteristics, values, and lifestyles are known as *psychographics*. Psychographics provide a richer portrait of the consumer than is possible through simple demographics or geographic variables (Zikmund and d'Amico, 2001). Thus psychographic segmentation is generally regarded to be more effective than either geographic or demographic segmentation, although information on psychographic variables is harder to obtain. Two of the most popular approaches to psychographic segmentation are VALS™ and AIO statements.

VALS™. The name VALS™ stands for values and lifestyles. VALS™ is a commercially available psychographic system developed by SRI Consulting and Business Intelligence, a research and consulting firm. The VALS™ system seeks to divide consumers into different psychographic segments based on two characteristics: (i) primary motivations that lead the consumer to seek a product, service, and/or experience, and (ii) the resources possessed by the consumer which enhance or constrain the expression of his or her primary motivation (Boone and Kurtz, 2001).

Consumer's purchase of products, services, and/or experiences reflects their preferences. It gives shape, substance, and satisfaction to their lives. Primary motivations are indicative of the meaningful core about the self or the world that governs the consumer's activities. The SRI Consulting and Business Intelligence website suggests that consumer's purchases can be inspired by one of three primary motivations: ideals, achievement, and self-expression (http://www.sric-bi.com/VALS/). Consumers primarily motivated by ideals are guided by knowledge and principles, whereas those motivated by achievement look for products and services that demonstrate success to their peers. Finally, consumers primarily motivated by self-expression desire social or physical activity, variety, and risk (Boone and Kurtz, 2001).

It should be noted that consumers' tendency to consume goods and services extends beyond simply their age, income, and education. Consumers' energy, self-confidence, intellectualism, novelty seeking, innovativeness, impulsiveness, leadership, and vanity also play a critical role in their tendency to consume. These personality traits together with key demographics of an individual determine the consumer's resources,

which can be either low or high (http://www.sric-bi.com/VALS/types.shtml).

On the basis of these three types of primary motivations (i.e., ideals, achievement, and self-expression), and two types of resources (i.e., low and high), VALS™ classifies consumers into eight different segments: innovators, thinkers, achievers, experiencers, believers, strivers, makers, and survivors. Each segment has a specific profile, behavioral and consumption pattern associated with it. For example, thinkers are comfortable, high-resource individuals who are motivated by ideals. They are mature and reflective, and value order, knowledge, and responsibility. Thinkers tend to be well educated, well informed, and seek out information in the decision-making process. Although their income allows them many choices, thinkers tend to be conservative, practical consumers, who look for durability, functionality, and value in the products they buy. The SRI Consulting and Business Intelligence website provides detailed information on the other VALS™ consumer segments (http://www.sric-bi.com/VALS/).

SRI Consulting and Business Intelligence have extended this basic approach to create several specialized psychographic systems. For example, GeoVALS™ estimates the percentage of each VALS™ segment in every residential zip code in the United States. Japan-VALS™ segments the Japanese marketplace on the basis of mind-set, emphasizing the identification of early adopters. iVALS™ focuses on Internet sites and users (Boone and Kurtz, 2001). Additional details on the VALS™ survey and other specialized VALS™ systems are available at the SRI Consulting and Business Intelligence website (http://www.sric-bi.com/VALS/). The website also lists several examples of the application of the VALS™ system ranging from identification of affluent early adopters by a European luxury automobile manufacturer to repositioning ubiquitous products in commodity categories by a leading US bank. VALS™ remains one of the most popular commercial psychographic segmentation systems.

AIO statements. AIO stands for activities, interests, and opinions. AIO statements are another popular approach to psychographic segmentation. Here data are collected from respondents related to the degree to which they agree/disagree with statements describing various activities, interests, and opinions (Boone and Kurtz, 2001). These data are used to develop lifestyle profiles of individuals, and group together individuals with similar lifestyle profiles. Marketers then develop separate marketing strategies for each lifestyle segment (Boone and Kurtz, 2001). Boone and Kurtz note that this approach is extensively used for a variety of products and markets ranging from beer to hospitals.

Product-related segmentation. Consumers can also be divided into segments based on their relationship to the product. This approach can take various forms, such as benefits that consumers seek from a product, usage rate for a product, and loyalty toward a product, among others (Boone and Kurtz, 2001).

Segmentation by benefits sought. People may buy the same product for very different reasons. Segmentation by benefits sought approach focuses on segmenting consumers based on the attributes they seek, and the benefit they expect to derive from a good or service. For example, Boone and Kurtz (2001) note that the same box of Arm & Hammer baking soda can end up serving as a baking ingredient, a refrigerator freshener, a toothpaste, or a deodorizer. Thus the baking soda consumers can be divided into various segments based on the reason they are buying the product, or the benefit they are seeking from it. Similarly, motorcycles, as a product, are perceived differently by different consumers. Motorcycles may be primarily perceived as a hobby or a "grown-up toy" in many affluent countries like the United States, whereas in developing countries like India they may be seen as an everyday mode of transportation for the individual or family. The same product would have to be marketed differently to consumers in these countries, or even consumers from different socioeconomic backgrounds within the same country.

Segmentation by usage rate. Consumers use a product or service at different rates. Some consumers tend to be heavy users and others

medium or light users of a product. It is often seen that a relatively small number of consumers account for a large proportion of a product's consumption. This is generally referred to as the *80/20 rule*, that is, 80% of the revenues produced by a product come from a small number (20%) of the customers (Boone and Kurtz, 2001). On the basis of the differences in the needs and preferences of heavy, medium, and light users, firms often market to these groups differently.

Frequent Flier Programs offered by airlines are a good example of this segmentation approach. It is common for airlines to offer priority boarding and baggage handling, free seat upgrades, discount tickets, vacation specials, and other kinds of perks to customers who fly more than a certain number of miles with them (e.g., Medallion members of the Delta Airline frequent flier program). Most airlines use a sliding scale of incentives; the more a consumer flies with them, the more perks are awarded to the consumer.

Segmentation by brand loyalty. Consumers can also be divided into segments based on the loyalty they feel toward a product or a brand. The airline frequent flier program mentioned above can also serve as an example of segmentation by brand loyalty. By offering attractive perks to heavy users and loyal customers, such programs seek to create incentives to develop brand loyalty to their product. Other industries like hotels, car rentals, and even movie rentals practice similar brand-loyalty-based segmentation (Boone and Kurtz, 2001). Research on CUSTOMER RELATIONSHIP MANAGEMENT, customer lifetime value analysis (*see* CUSTOMER LIFETIME VALUE (CLV)), and database marketing (*see* DATABASE MINING AND MARKETING) shows that segmenting markets based on brand loyalty and usage rates can significantly increase product margins and even firm profitability.

Some marketing scholars (e.g., Zikmund and d'Amico, 2001) categorize product-related segmentation differently by making a distinction between behavior-pattern-based segmentation (e.g., type of store, shopping frequency, media habits, etc.), consumption-pattern-based segmentation (e.g., usage rate, brand loyalty, etc.), and consumer-predisposition-based segmentation (e.g., benefits sought, product knowledge, etc.). A more nuanced treatment of these distinctions is beyond the scope of the present discussion due to space constraints. However, the basic principles discussed above apply in these cases as well. Next, we discuss the bases for segmenting customers in business markets.

Bases for Segmenting Business Markets

Wind and Cardozo (1974) suggest that the bases for segmenting business markets can broadly be classified into two major categories: macrosegmentation and microsegmentation.

Macrolevel bases. Macrosegmentation centers around dividing the market based on the characteristics of the buying organization (e.g., size, location, industry classification, and organizational structure), and the buying situation (e.g., straight rebuy, modified rebuy, and new task buying).

- *Characteristics of buying organizations* (e.g., size, location, user rate, structure of procurement, etc.). Dividing business markets based on the size of the buying organization is a commonly used segmentation approach. Large organizations often display unique requirements, and respond to marketing mix stimuli differently than smaller organizations (Bellizzi, 1981). Organizations that are heavy users of a firm's product may have needs different than those of light users. For example, businesses that are heavy users may place greater value on technical support or reliable delivery. Similarly, firms having centralized procurement functions may behave differently and emphasize different criteria than firms where the procurement function is decentralized (Hutt and Speh, 2010). Nowadays, the major differences can be seen in firms with primarily online procurement processes versus those with primarily offline procurement processes. Sometimes location of the buying firm can affect issues like the cost, delivery, and level of regulation involved. All of the above factors can influence the buying behavior of business customers, and be

used as a basis for differentiating marketing strategies.

- *Product/Service application* (e.g., industrial classification, end market served, value in use, etc.). Industrial goods can be used in many different ways (e.g., industrial gases). Often the applications and the needs of the customers can vary based on the industry of the buying organization. Hence business marketers can segment industrial markets based on industry classification, end-use application, end market served, or the product's economic value to the user relative to a specific alternative in a specific application (Hutt and Speh 2010). For example, Hutt and Speh note that Intel's microprocessors can be used in many different applications and industries ranging from household appliances, cell phones, computers, to toys. The marketing strategies employed by Intel for each of these industries would differ considerably based on end-use application and industry-specific requirements and needs.
- *Characteristics of purchasing situation* (e.g., type of buying situation, stage of purchase decision process, etc.). Information and customer service needs of a first-time buyer (i.e., new task buy) are likely to be very different from those of a repeat buyer (i.e., straight rebuy or modified rebuy). Thus the position of the buying firm in the purchase decision process (i.e., early or late) or the buying situation can dictate the marketing strategies employed. For example, vendor comparison/evaluation may play a much more prominent role in a new task buy or modified rebuy situations compared to a straight rebuy situation. Thus marketers can gain sales by designing appropriate strategies for customers in the new task or modified rebuy situations. Conversely, failure to account for differences in customers based on purchase situations can leave marketers vulnerable to competitive action.

Microlevel bases. Microsegmentation focuses on the decision-making units within each macrosegment and requires greater marketing knowledge (e.g., key buying decision criteria, attitude toward the vendors, and perceived importance of the purchase) (Wind and Cardozo, 1974).

- *Key criteria* (e.g., vendor reliability, price, technical support, delivery, etc.). For some business products, markets can be divided on the basis of key criteria that customers consider most important for the purchase decision (Hutt and Speh, 2010). For example, based on the value proposition sought by customers (*see* A FRAMEWORK FOR CREATING VALUE PROPOSITIONS; VALUE PROPOSITION), the chemical giant Dow Corning segments its markets into innovation-focused customers, customers in fast-growing markets, and customers in highly competitive markets (Hutt and Speh, 2010). Similarly, Rangan, Moriarity, and Swartz (1992) note that in mature industrial markets, customers could be segmented into four categories based on price-service trade-offs: programmed buyers (i.e., buyers insensitive to price or service), relationship buyers (i.e., buyers that value partnerships and do not push for price or service concessions), transaction buyers (i.e., buyer that actively consider price vs. service trade-offs), and bargain hunters (i.e., buyers who are very sensitive to any price or service change).
- *Purchasing strategies* (e.g., single sourcing vs. multiple sourcing). Some business markets can be segmented on the basis of the buyer's purchasing strategies. Some buyers purchase from a single supplier, or a very small number of suppliers, giving each a healthy share of their business in exchange for a stronger relationship and preferred treatment (Hutt and Speh, 2010). Other buyers tend to employ a larger number of suppliers in an effort to reduce their risk, prevent any disruption in their supplies, or even maintain a stronger position relative to their suppliers.
- *Structure of the decision-making unit* (e.g., purchase manager vs. buying center, etc.). Structure of the decision-making unit at the customer end (or buying center) can also provide a basis for segmenting business markets. Depending upon which participants are involved in (or exert

influence over) the purchasing process (e.g., purchasing manager, engineering, senior management, etc.), marketers can fine-tune their presentation or marketing mix in each segment to take into account the dynamics of the decision-making units. Sometimes this approach also takes into account the personal characteristics (i.e., age, education, decision-style, and risk propensity) of the key decision makers at the customer end.

- *Organizational innovativeness* (e.g., innovator, follower, etc.). Robertson and Wind (1980) note that firm psychographics and demographics can be used to predict the adoption of new products by organizations. For example, some organizations are resistant to change, while others are more innovative and have a desire to excel. Robertson and Wind suggest that the latter are more likely to purchase new industrial products than the former. Marketers can use such organizational innovativeness variables to identify segments of business markets to target first when introducing new products (Hutt and Speh, 2010).

Targeting

Owing to resource constraints, firms are often unable to pursue all the segments they identify in a given market. Typically, firms focus their marketing resources on satisfying a smaller number of segments in the market (Zikmund and d'Amico, 2001). This process of allocating resources effectively by focusing marketing efforts on a selected part of the total market is called *targeting* (Zikmund and d'Amico, 2001). By focusing their limited resources on the targeted segments, firms are able to compete effectively in those segments by developing a distinct value proposition better than their competitors (Hutt and Speh 2010). The targeting strategies that firms follow can broadly be classified into four major categories: undifferentiated marketing, differentiated marketing, concentrated marketing, and micromarketing (Best 2004; Boone and Kurtz 2001; Porter 1985; also *see* COMPETITIVE ADVANTAGE).

Undifferentiated marketing. In undifferentiated marketing, a firm produces only one product or product line, and promotes it to all the customers with a single marketing mix (Boone and Kurtz, 2001; Porter, 1985). Such an approach is also referred to as *mass marketing*. An undifferentiated marketing approach is seen when differences in customer needs and/or demographics are either small or indistinct (Best 2004). This targeting strategy presents a generic business strategy with generic positioning, built around a core product (Best, 2004).

While such a strategy might be efficient from a production standpoint, it leaves firms vulnerable to competitive action (Boone and Kurtz, 2001). Firms following an undifferentiated targeting approach can suffer if competitors are able to successfully differentiate the market. That is, competitors enter the market with specialized products and marketing mix that are able to better meet the needs and preferences of smaller segments of consumers within the overall market (*see* COMPETITIVE ADVANTAGE). Although less frequently used now, a prime example of undifferentiated strategy is the marketing of the Ford Model T. Henry Ford built the Model T and sold it at one price to everyone. He was famously known to have said that he was willing to give the customers any color car they wanted – as long as it was black (Boone and Kurtz, 2001).

Differentiated marketing. Under this strategy, firms promote numerous products with differing marketing mixes designed to satisfy customers in smaller segments (Boone and Kurtz, 2001). In each segment, the firm selects one or more dimensions that the customers value and uniquely positions itself to meet those needs – while charging a premium price for such uniqueness (*see* POINT OF DIFFERENCE AND PRODUCT DIFFERENTIATION; COMPETITIVE ADVANTAGE; PRODUCT POSITIONING). Porter (1985) suggests that several sources can provide the bases for differentiation, such as the product itself, the method of delivery/distribution, the marketing approach, and so on. Compared to an undifferentiated approach, a differentiated strategy allows a firm to increase not only its total sales (Boone and Kurtz, 2001), but also its average margin per sale (Best, 2004), by providing increased satisfaction to each of the targeted segments.

A differentiated strategy increases marketing costs. However, despite increasing costs, firms are forced to practice differentiation in order to remain competitive (Boone and Kurtz, 2001). Porter (1985) suggests that in order to remain competitive and counter increasing costs associated with a differentiation strategy, firms must strive for cost parity with their competitors by reducing costs in areas that do not affect differentiation.

Many products and services follow a differentiation strategy. For example, most major airlines offer economy class, business class, and first class seating based on how much value passengers place on their comfort and service while traveling. Similarly most automobile companies offer a variety of cars ranging from sedans, to convertibles, to minivans, and so on, based on the needs and preferences of the various segments in the market.

Concentrated marketing. Rather than trying to market its product or service to multiple segments, a firm may choose to use its marketing resources to profitably satisfy customers in a single market. Such a targeting strategy is known as *concentrated*, *focus*, or *niche marketing* (Boone and Kurtz, 2001; *see* COMPETITIVE ADVANTAGE). By deliberately electing to narrow the scope of its activities, a firm seeks to achieve competitive advantage in the niche segment targeted, even though it may not enjoy a competitive advantage overall (Porter, 1985). A concentrated marketing strategy works well when the targeted niche segment has needs or preferences that are poorly served by more broadly targeted competitors. Under such conditions, a firm can exploit these differences and achieve competitive advantage by dedicating itself to serving the niche segment(s) exclusively (Porter, 1985).

A focus/niche targeting strategy generally appeals to small firms that lack the financial resources of its competitors, or to companies that offer highly specialized goods and services (Boone and Kurtz, 2001). Sub-Zero Refrigerators are a good example of successful niche segmentation. Sub-Zero specializes in very expensive, built-in refrigerators that start at $3500 per unit. While Sub-Zero has less than a 2% share of the overall refrigerator market in the United States; it commands a 70% market share of the "Super Premium" (i.e., high end) segment – a niche segment within the refrigerator market (Best, 2004).

Micromarketing. Micromarketing, also known as *mass customization*, refers to a focus even narrower than concentrated or niche marketing. It represents the polar opposite of an undifferentiated strategy. Micromarketing involves targeting potential customers at a very basic level, that is, postal code, occupation, household, and so on, and ultimately the individuals themselves (Boone and Kurtz, 2001; *see* MASS CUSTOMIZATION STRATEGIES). Under this targeting approach, the smallest group of unique customers is treated as a niche segment, and the firm develops a customized segment strategy for it, even if the segment consists of a single individual (Best 2004).

Increasing use of database marketing (*see* DATABASE MINING AND MARKETING) and internet marketing (*see* E-COMMERCE AND INTERNET MARKETING) allows the tracking of specific demographic, personal, and web browsing information, making it easier for marketers to engage in effective micromarketing (Boone and Kurtz, 2001). For example, Amazon.com, the giant online book and music retailer, provides its customers with individualized recommendations based on his/her web browsing patterns and past purchase behavior.

Conclusion

The concepts of segmentation and targeting lie at the heart of marketing. Segmentation is the process of dividing heterogeneous markets into homogeneous subgroups, which display similarities in needs, preferences, and/or behaviors. Segmentation allows firms to develop a marketing mix better suited to the customers in each segment. Four criteria are used to determine effective segmentation: measurability, accessibility, substantiality, and responsiveness of the segment(s) selected. However, due to resource constraints, firms are often unable to pursue all the segments they identify in a market. They have to select or target the segments they would like to pursue.

The preceding discussion briefly outlines the bases commonly used to segment consumer and

business markets. In consumer segmentation, we discussed geographic, demographic, psychographic, and product-related approaches to segmenting markets. In business segmentation, we covered macrolevel bases of segmentation (i.e., characteristics of the buying organization, product/service application, and characteristics of the purchase situation) as well as microlevel bases of segmenting business markets (i.e., key criteria, purchasing strategies, structure of decision-making units, and organizational innovativeness). We conclude by describing four popular targeting strategies: undifferentiated marketing, differentiated marketing, macromarketing, and micromarketing.

Owing to space constraints, the discussion here cannot exhaustively cover all the issues pertinent to market segmentation and targeting. Readers may wish to refer to the articles on positioning (*see* POSITIONING ANALYSIS AND STRATEGIES), pricing (*see* PRICING STRATEGY), and forecasting (*see* NEW-PRODUCT FORECASTING), for additional details on related topics.

Bibliography

Best, R.J. (2004) *Market-Based Management: Strategies for Growing Customer Value and Profitability*, 3rd edn, Prentice Hall, Upper Saddle River, NJ.

Bellizzi, J.A. (1981) Organizational size and buying influences. *Industrial Marketing Management*, **10**, 17–21.

Boone, L.E. and Kurtz, D.L. (2001) *Contemporary Marketing*, 10th edn, Harcourt College Publishers, New York, NY.

Hutt, M.D. and Speh, T.W. (2010) *Business Marketing Management: B2B*, 10th edn, South-Western Cengage Learning, Mason, OH.

Porter, M.E. (1985) *Competitive Advantage: Creating and Sustaining Superior Performance*, The Free Press, New York, NY.

Rangan, V.K., Moriarity, R.T., and Swartz, G.S. (1992) Segmenting customers in mature industrial markets. *Journal of Marketing*, **56**, 72–82.

Robertson, T.S. and Wind, Y. (1980) Organizational psychographics and innovativeness. *Journal of Consumer Research*, **7**, 24–31.

Wind, Y. and Cardozo, R.N. (1974) Industrial market segmentation. *Industrial Marketing Management*, **3**, 155.

Zikmund, W.G. and d'Amico, M. (2001) *Marketing: Creating and Keeping Customers in an E-Commerce World*, 7th edn, South-Western College Publishing, Cincinnati, OH.

market share

Derek N. Hassay

Perhaps the most commonly referenced marketing metric (*see* MARKETING METRICS) Market share is defined as *the proportion of a market segment* (*see* MARKET SEGMENTATION AND TARGETING) *served by a company*. Market share can be reported as a percentage of revenue or unit sales volume according to the following formula:

$$\text{Market share}(\%) = \frac{\text{Company (unit or \$) sales}}{\text{Total industry (unit or \$) sales}} \times 100\% \quad (1)$$

An especially powerful metric, market share reflects both a company's performance relative to its *competitors* (*see* COMPETITIVE ANALYSIS) and the success of its MARKETING STRATEGY. Though simple to calculate, in practice, determining *competitive sales* (*see* COMPETITOR ANALYSIS) or consumer purchases for a given PRODUCT CATEGORY in a specific market can be a significant challenge. For instance, capturing sales information for a geographic market segment is far simpler than for a demographic (e.g., young males, low income) or psychographic (e.g., recreational athletes) segment. Moreover, the actual definition of a market is subjective and thus by narrowly defining its served market, a firm can report a larger market share and thus enhance its *market position* (*see* POSITIONING ANALYSIS AND STRATEGIES). For instance, a manufacturer of automobiles might suggest that it is the sales leader in the Northwest, where the Northwest has been narrowly defined as Washington State.

Market share is commonly cited as a key determinant of profitability, as firms with larger shares are generally thought to be more profitable than those with smaller shares in the same market because of economies of scale, reduced cost of marketing relative to sales, and higher prices (Buzzell, Gale and Sultan, 1975). The Profit Impact of Marketing Strategy (PIMS: Buzzell and Gale, 1987) studies, conducted by the Strategic Planning Institute, identify market share and relative market share (the proportion of a firm's share relative to its largest competitor) as two of 18 key strategic variables influencing business success or profitability. The often cited PIMS studies, highly regarded because of the volume of data they have captured on strategic business variables since the early 1960s (Strategic Planning Institute), indicated that market position factors such as market share have the most significant impact on profitability.

Despite the volume of articles, textbooks, and frameworks reporting market share as a key determinant of profitability, having greater market share is not necessarily synonymous with increased profitability. Schwalbach (1991), for example, demonstrated that the proposed linear relationship between share and profitability is flawed and argued that the true relationship between share and profitability is much more complex. In fact, it was suggested that while *correlated* (*see* PRODUCT MOMENT CORRELATION), the *causal relationship* (*see* CONCEPT OF CAUSALITY AND CONDITIONS FOR CAUSALITY) between share and profitability may actually be reversed with higher profits enabling greater market share growth (Davidson, 1985). Subsequent research efforts revealed that the theoretical underpinnings of the share–profitability relationship – the experience curve – did not have the pronounced impact on profitability as originally thought nor did it hold across different types of industries (Davidson, 1985). This information prompted the Boston Consulting Group to produce a new BCG Matrix (Hax and Majluf, 1983), as its original portfolio planning tool based on relative market share and economies of scale was inaccurate. The new BCG matrix examined the relationship between ROI and market share in four unique industry contexts (Hax and Majluf, 1983).

At the very least, profitability is likely to suffer in the short term when a firm pursues market share growth because competitor-oriented objectives such as this typically require increased marketing expenditures. And in those cases where a firm is close to capacity or has to expand *distribution* (*see* MARKETING CHANNEL STRATEGY) beyond regional or national markets, any increase in market share will require a concomitant investment in production, inventory, and distribution capabilities.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Recent research calls for firms to become more customer than competitor-centric and to manage *customer relationships* (*see* CUSTOMER RELATIONSHIP MANAGEMENT) for long-term profitability. Reinartz and Kumar (2002) found that some customers actually have a negative impact on profitability, while Zeithaml, Rust and Lemon (2001) cautioned firms against using profitable customers to subsidize the service of unprofitable ones. Sharma (2008) argued that rather than spending valuable marketing efforts on unprofitable customers, firms may need to outsource or even fire some customers, a position which suggests that, in these instances, a lower market share is likely to lead to increased profitability.

Bibliography

Buzzell, R.D. and Gale, B.T. (1987) *The PIMS Principles: Linking Strategy to Performance*, Free Press, New York.

Buzzell, R.D., Gale, B.T., and Sultan, R.G.M. (1975) Market share – a key to profitability. *Harvard Business Review*, **53**, 97–106.

Davidson, K. (1985) Strategic investment theories. *The Journal of Business Strategy*, **6** (1), 16–28.

Hax, A.C. and Majluf, N.S. (1983) The use of the growth-share matrix in strategic planning. *Interfaces*, **13** (1), 46–50.

Reinartz, W. and Kumar, V. (2002) The mismanagement of customer loyalty. *Harvard Business Review*, **80** (7), 86–94.

Schwalbach, J. (1991) Profitability and market share: a reflection on the functional relationship. *Strategic Management Journal*, **12** (4), 299–306.

Sharma, A. (2008) Improving customer service and profitability through customer intervention in service relationships. *Journal of Relationship Marketing*, **7** (4), 324–340.

Zeithaml, V.A., Rust, R.T., and Lemon, K.N. (2001) The customer pyramid: creating and serving profitable customers. *California Management Review*, **43** (4), 118–142.

market/industry structure

Joseph A. Cote

Economics categorizes markets as perfect competition, monopolistic competition, oligopoly, or monopoly – with a focus on policy implications. Marketing differs, emphasizing the managerial implications of market structure. Marketing defines market/industry structure as the characteristics of a competitive space (*see* COMPETITIVE ANALYSIS) that affect MARKETING STRATEGY and tactics. The effects of market/industry structure can be examined using a wide variety of frameworks including Porter's model, the marketing environment, critical success factors (*see* SUCCESS FACTORS FOR NEW-PRODUCT DEVELOPMENT), the resource-based view, stakeholder relationships, network effects, value chain analysis, product life cycle (*see* STAGES OF THE PRODUCT LIFE CYCLE; PRODUCT LIFE CYCLE), and prototypical business models (*see* MARKETING STRATEGY MODELS). The more prominent frameworks are summarized below.

First Step

The market/industry must be defined before conducting a market/industry structure analysis (*see* MARKET ORIENTATION). NAICS or SIC classification codes can be used, but a rigorous analysis should consider the levels of competition. Competition can occur within a PRODUCT CATEGORY (e.g., beer), across closely related product categories (e.g., wine and liquor), or across any relevant product category (e.g., beverages). Knowing the alternatives considered by consumers (*see* CUSTOMER ANALYSIS) allows you to more accurately define the competitive space. For example, cruise companies not only compete among themselves but also with a host of other vacation and "staycation" options.

Porter's Model

Porter's model is often used to evaluate the overall attractiveness of an industry (Porter, 2008; Porter, 1979). He examines industry attractiveness using five forces: existing competition (*see* COMPETITOR ANALYSIS), buyers (*see* MARKETING CHANNEL STRATEGY), suppliers (*see* SUPPLY CHAIN MANAGEMENT STRATEGY), substitutes, and potential competitors. Porter argues that, "Industry structure, as manifested in the strength of the five competitive forces, determines the industry's long-run profit potential because it determines how the economic value created by the industry is divided . . . " (Porter, 2008, p 86). For example, high economies of scale can prevent new competitors from entering a market, lowering competitive intensity. Conversely, DEMAND ELASTICITY among buyers will increase competitive intensity. Understanding the five forces also provides insights into prices, costs, necessary investments, and potential shifts in the industry.

The Marketing Environment

A marketing environment analysis is used to understand the underlying drivers that determine success or limit a firm's actions (see Figure 1). In addition to suppliers, buyers, and competitors, the marketing environment considers macrolevel conditions and trends related to economic, political/legal, technological, and social/cultural/demographic (*see* SOCIETY, CULTURE, AND GLOBAL CONSUMER CULTURE) issues. For example, during the 2008–2009 economic crisis, auto makers had to significantly change how they did business since consumers were either unable or unwilling to obtain loans to purchase new vehicles. Similarly, demographics trends (aging population and increasing incidence of obesity) have significant implications for the health-care industry and proposed regulations of greenhouse gases may significantly affect power generation and energy companies.

Critical Success Factors and Key Resources

Leidecker and Bruno define critical success factors (CSFs) as "those characteristics, conditions, or variables that when properly sustained, maintained, or managed can have a significant impact on the success of a firm competing in a particular industry." (Leidecker and Bruno, 1984, p 24). For example, customer retention (*see* CUSTOMER LIFETIME

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

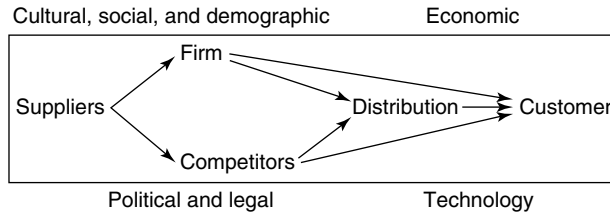


Figure 1 The marketing environment.

VALUE (CLV); CUSTOMER RELATIONSHIP MANAGEMENT) is often a CSF for monthly subscription services such as cell phones and health clubs. Once the critical success factors have been identified, firms can then develop strategies and key resources (*see* CORE COMPETENCIES) to address these factors, thus creating COMPETITIVE ADVANTAGE (*see* COMPETITIVE ADVANTAGE: ITS SOURCES AND THE SEARCH FOR VALUE) (per the resource-based view, see Barney, 1991).

Stakeholder Analysis

Freeman 1984 suggests that an organization's success depends on its relationships with employees, customers, value chain partners, financiers (investors, banks, etc), and the public (communities, government, etc). Understanding how each stakeholder affects the organization helps management design strategy to best leverage these relationships. Stakeholders often provide the key resources necessary for competitive advantage. For example, Apple's ability to work with content providers on digital rights management allowed them to create iTunes and a dominant position in music distribution. Stakeholder analysis is often used to look beyond the obvious players to understand how others affect an industry (Fassin, 2009).

Network Effects

Increasingly, industries are defined by networks rather than individual companies or products. Network relationships provide access to information, markets, technologies, etc., and the pattern of network relationships can affect performance (Gulati, Nohria, and Zaheer, 2000). For example, the iPhone relies on a network of independent developers to create applications. The types and number of applications help create competitive advantage for the iPhone. Networks can also be product based (Lee and O'Connor 2003). For example, digital cameras became more useful when flash memory, desktop printers, image processing software, and CCD technology improved.

Bibliography

Barney, J. (1991) Firm resources and sustained competitive advantage. *Journal of Management*, **17**, 99–120.

Fassin, Y. (2009) The stakeholder model refined. *Journal of Business Ethics*, **84**, 113–135.

Freeman, R.E. (1984) *Strategic Management: A Stakeholder Approach*, Pitman Publishing Inc, Boston.

Gulati, R., Nohria, N., and Zaheer, A. (2000) Strategic networks. *Strategic Management Journal*, **2**, 203–215.

Lee, Y. and O'Connor, G.C. (2003) New product launch strategy for network effects products. *Journal of the Academy of Marketing Science*, **31**, 241–255.

Leidecker, J.K. and Bruno, A.V. (1984) Identifying and using critical success factors. *Long Range Planning*, **17**, 23–32.

Porter, M.E. (1979) How competitive forces shape strategy. *Harvard Business Review*, **47**, 137–145.

Porter, M.E. (2008) The five competitive forces that shape strategy. *Harvard Business Review*, **86**, 78–93.

marketing audit

Bruce H. Clark

The term *marketing audit* is generally thought to have been introduced into the marketing literature in an American Management Association report (American Management Association, 1959). For a thorough history of the concept and its development, see Rothe, Harvey, and Jackson (1997). The use of the term *audit* is an explicit analogy to the accounting audits performed for companies on a regular basis.

Definition and Purposes

The most widely used definition of the construct is provided by Kotler, Gregor, and Rodgers (1977) in an influential *Sloan Management Review* article. They define a marketing audit as follows:

> A marketing audit is a comprehensive, systematic, independent, and periodic examination of a company's—or business unit's—marketing environment, objectives, strategies, and activities with a view to determining problem areas and opportunities and recommending a plan of action to improve the company's marketing performance (p. 27).

Kotler, Gregor, and Rodgers (1977) further distinguish between a *horizontal audit* covering the full scope of a company's marketing activities, and a *vertical audit* that examines a particular marketing function in depth (e.g., a sales force audit or an advertising audit).

As the definition suggests, audits are both backward-looking and forward-looking. They seek to evaluate past and current marketing assets (*see* MARKET-BASED ASSETS) and activities in the context of market conditions, but use the resulting analysis to aid the firm in planning. Some writers refer to auditing explicitly as a control function in this regard (Brownlie, 1996). Mokwa (1986) also notes that audits can perform an important educational function for the firm in terms of improving awareness and comprehension of marketing issues.

Elements of Audits

Unlike an accounting audit, there are no formal standards for conducting marketing audits. Both Mokwa (1986) and Brownlie (1996) argue that this is a truth rather than a limitation: audits invariably must be customized to the firm being audited. That said, most discussions suggest an audit should cover both an analysis of the external situation facing the unit and a thorough review of internal marketing goals, strategies, capabilities, processes, and systems. In this sense, a formal situation analysis or SWOT ANALYSIS is an important part of the audit. Audits differ from SWOT analysis, where the "comprehensive, systematic, independent, and periodic review" indicated in the definition generates a deeper, more objective (and iterative) learning, and planning process. The literature contains several comprehensive frameworks that cover likely areas of interest to auditors (*see* Kotler, Gregor, and Rodgers, 1977; Mokwa, 1986; Berry, Conant, and Parasuraman, 1991; Brownlie, 1996; Wilson, 2002).

The Audit Process

Auditors use an audit framework to conduct a structured consulting project that identifies ways in which the audited firm can improve its marketing. The audit may be conducted by either internal or external auditors, but auditors must be independent of the audited unit in either case. Most authors argue that conducting regular, periodic audits maximizes the value of the process. For a given firm, having a common framework applied over time will aid in developing more long-term learning. Regular audits also avoid the fear employees sometimes have that an ad hoc audit is a prelude to unwelcome changes. Brownlie (1996) covers audit process issues particularly well. Enis and Garfein (1992) also outline a computerized approach to auditing.

Empirical Evidence

It is difficult to tell how widespread marketing audits are, partly because many firms may engage in elements of an audit (e.g., situation analysis) without calling the process an "audit." Trade coverage of audits suggests service applications are particularly common (*see* Berry, Conant, and

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Parasuraman, 1991 for good coverage of the elements of a service audit). Despite a compelling logic to the process, there is also little evidence beyond case studies regarding the benefits of conducting audits. One exception is Taghian and Shaw (2008), who find a positive correlation between conducting marketing audits and (self-reported) change in market share and financial performance.

Relationship to Other Constructs

The marketing audit is typically seen as one way of evaluating the health of a firm's marketing practices. In this sense, it is strongly related to broader research on marketing performance measurement (Bonoma and Clark, 1988), marketing control (Jaworski, 1988), MARKET ORIENTATION, and the more recent movement toward marketing accountability and MARKETING METRICS.

Bibliography

American Management Association (1959) in *Analyzing and Improving Marketing Performance: Marketing Audits in Theory and Practice* (eds A. Newgarden and E.R. Bailey) American Management Association, New York Report No. 32.

Berry, L.L., Conant, J.S., and Parasuraman, A. (1991) A framework for conducting a services marketing audit. *Journal of the Academy of Marketing Science*, **19** (3), 255–268.

Bonoma, T.V. and Clark, B.H. (1988) *Marketing Performance Assessment*, Harvard Business School Press, Boston.

Brownlie, D. (1996) The conduct of marketing audits: a critical review and commentary. *Industrial Marketing Management*, **25**, 11–22.

Enis, B.M. and Garfein, S.J. (1992) The computer-driven marketing audit: an interactive approach to strategy development. *Journal of Management Inquiry*, **1** (4), 306–317.

Jaworski, B.J. (1988) Toward a theory of marketing control: environmental context, control types, and consequences. *Journal of Marketing*, **52** (3), 23–39.

Kotler, P., Gregor, W., and Rodgers, W. (1977) The marketing audit comes of age. *Sloan Management Review*, **18** (2), 25–43.

Mokwa, M.P. (1986) The strategic marketing audit: an adoption utilization perspective. *Journal of Business Strategy*, **6** (4), 88–95.

Rothe, J.T., Harvey, M.G., and Jackson, C.E. (1997) The marketing audit: five decades later. *Journal of Marketing Theory and Practice*, **5** (3), 1–16.

Taghian, M. and Shaw, R.N. (2008) The marketing audit and organizational performance: an empirical profiling. *Journal of Marketing Theory and Practice*, **16** (4), 341–349.

Wilson, A. (2002) *The Marketing Audit Handbook: Tools, Techniques & Checklists to Exploit your Marketing Resources*, Kogan Page Publishers, London.

marketing costs

Mukesh Bhargava

Almost a century ago, Shaw (1912) explored the problems in distribution focusing on reducing marketing costs and improving the productivity. The themes of cutting marketing costs and increasing efficiency explored in that paper still resonate today. Marketing costs account for the largest part of the consumer dollar. The prefix "cutting" features in most current articles on marketing costs. Further, a large number of articles link lack of profitability to escalating marketing costs. We need to build the perception that marketing costs produce value, and should therefore be viewed as investments in building the assets of a firm (*see* MARKET-BASED ASSETS).

Marketing Costs and Tactics

Most marketing textbooks discuss cost under a section on Marketing Arithmetic covering topics such as margins, fixed and variable costs, and breakeven analysis. Marketing costs are mentioned, usually qualitatively, while detailing the marketing mix decisions. Some of the examples are costs associated with the following:

1. Product quality, packaging, logo and brand development, and brand extensions.
2. Price decisions covering fixed and variable costs, breakeven analysis, and the demand (*see* DEMAND ELASTICITY).
3. Distribution requiring an analysis of transportation, inventory carrying, distribution margins, incentives to "push" products including allowances to trade.
4. Communication budgets including allocation by promotion mix and media selection, and agency compensation (*see* COMMUNICATIONS BUDGETING).

Most material does not provide any detail about the amounts as the figures vary by product, product lifecycle, and country. However, in some industries such as pharmaceuticals and healthcare, and government and nonprofit organizations, marketing costs are often portrayed as being wasteful (Gagnon and Lexchin, 2008).

Marketing Costs and Strategy

Strategically, lower costs provide competitive advantage. The costs can be lower because of economies of scale, where the volume provides a lower unit cost per item. The other methods of decreasing costs include lowering costs of production through a learning curve. Additionally, economies of scope also can lead to lower costs. These include cost advantage in brand extensions, distribution coverage as well as the use of marketing knowledge and processes for new-product development and commercialization. Cost analysis provides an indication of the "strategic group" that the firm will be competing against. All of these factors can lead to greater efficiencies, assuming that the company is doing "things" right. Unfortunately, while cost-based metrics help us to improve the marketing efficiency, it is the issues of being effective that really need to be addressed.

Marketing Costs as Investments

In the last decade, there has been a renewed interest in computing the return on marketing investments, that is, trace the costs and the benefits. The presence of better databases allowed the estimation of cost of new-customer acquisition and the lifetime value of a customer. However, generally acceptable practices for relating costs and results, the foundation to the application of tools from accounting and finance, have yet to emerge (Farris *et al.*, 2006). (*see* MARKETING METRICS).

There are numerous complicating factors that introduce complexity in developing the "true" returns of marketing investments. Marketing costs are expensed over a time period, whereas there is adequate research that shows that the outcomes (e.g., new customers, creating a brand name, better relationship with channel members) lead to an increased cash flow in the future. Other complicating factors are as follows:

1. Marketing activities are dispersed both within and outside the functional unit.
2. Manufacturing costs are allocated over products, whereas the marketing costs allocation has multiple, conflicting basis such as products, territories, channels, size of orders, and even sales force.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

3. The outcomes are influenced by competitive actions and other exogenous variables over which the company and the manager do not have direct control.
4. There are numerous alternatives for product development, pricing, communication, and distribution, which merge online and offline tactics.

The analysis of the cost of marketing can be done at the national, industry, firm, brand, channel, advertising campaign, package sizes, and customer levels (Doyle, 1998). Aggregation of data by time period and level of analysis often masks vital information necessary for any meaningful analysis. The crucial problem is that relevant cost data that companies are willing to share is extremely limited. As a result, there is a paucity of academic work on net results of marketing activities (net margins, cash flow, etc.).

The silver lining in the appropriate use of costs and returns as managerial tools in marketing is the development of fine-grained data for each customer and product (stock keeping unit (SKU)) and time periods. Given these developments, it is possible to conduct experiments that provide a better measure of marketing costs and returns.

Bibliography

Doyle, D. (1998) *Adding Value to Marketing: The Role of Marketing in Today's Profit Driven Organization*, Kogan Page, London.

Farris, P.W., Bendle, N.T., Pfeifer, P.E., and Reibstein, D.J. (2006) *Marketing Metrics: 50+ Metrics Every Executive Should Master*, Wharton School Publishing, Upper Saddle River.

Gagnon, M.-A. and Lexchin, J. (2008) The cost of pushing pills: a new estimate of pharmaceutical promotion expenditures in the united states. *PLoS Medicine*, **5** (1), e1. doi: 10.1371/journal.pmed.0050001.

Shaw, A.W. (1912) Some problems in marketing distribution. *Quarterly Journal of Economics*, **26**, 706–765.

marketing channel strategy

Anne T. Coughlan

Introduction

Marketing channel strategy – one of the "Four P's" of the marketing mix ("place," along with product, price, and promotion (*see* MARKETING MIX)) – comprises the set of decisions made, and structures created, that help move a manufacturer's product or service from its place of manufacture to the ultimate end-user. Unlike the other three "P's," it inherently involves collaboration and interaction among multiple corporate entities, all in the pursuit of a single goal: profitable end-user sales. The strength of partnering with other companies or entities is the access it affords to partners' capabilities. The corresponding challenge is the structuring and ongoing management of a complex set of activities, undertaken across corporate borders, without complete control. This article lays out the key elements of effective channel design and channel management.

Marketing Channel Definition and Elements

A "*marketing channel*" (also called a "*distribution channel*" or "*route to market*") is defined as follows:

A marketing channel is a set of ***interdependent*** organizations involved in the ***process*** of making a product or service ***available for use or consumption***.[1]

This definition highlights three concepts important in managing a marketing channel. Consider the italicized terms in the definition, in reverse order.

First, a channel's very purpose is to make its product or service *available for use or consumption*. This means that the primary focus is always on end-users and their satisfaction. As a marketing concept, this hardly seems earth-shattering; after all, do we not start every analysis with the consumer or end-user? Sadly, this is often *not* the case when managers consider their channel strategies. For example, a manufacturer selling to the ultimate end-user through a retailer may erroneously consider the retailer a "customer." In fact, the retailer is not the manufacturer's customer, but a *partner* in *jointly* serving the ultimate end-user (who is the true consumer).

An easy way to verify the identity of the end-user is to ask the question, "Who pays the channel's bills?" Only ultimate end-users inject fresh money into the channel and thus pay for all the costs of running the channel, as well as any profits above and beyond those enjoyed by all channel members.

Secondly, running a marketing channel is a *process*, rather than an "event" happening at the moment of sale. Distribution involves the cumulative investment of time and effort, usually by multiple channel members, engaged in different costly and value-added activities, which together culminate in a sale. These activities involve coordination among multiple partners, communication forward and backward through the channel, and the participation of payment facilitators and financing entities in addition to those partners handling the physical inventory itself.

Thirdly, keeping a marketing channel running well means recognizing and managing the *interdependence* of channel partners. Different channel members have responsibility for different functions and activities in the marketing channel, and therefore each is dependent on the good offices of all the other channel partners. Given this interdependence, attention must be paid to the possibility of channel conflicts that could disturb channel partners' incentives to continue performing their assigned roles. Should any such conflicts arise, the judicious application of levers of channel power can help to keep conflict from disrupting the channel's operation.

This article uses these three elements as core concepts in describing channel design and channel management and implementation in more detail below, after discussing the inherently *strategic* nature of channel design and management.

Why Marketing Channel Design and Management is Strategic

The design and management of a well-working marketing channel not only takes considerable effort and thought, but it affects the productivity

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

of the rest of the company's marketing efforts as well. An excellent new product idea may enjoy a certain amount of market potential. However, unless the product is distributed appropriately and supported by the efforts of intermediaries, it may fail miserably in the marketplace. Conversely, a well-orchestrated distribution effort can significantly increase the productivity of marketing investments in product design.

Further, the design and execution of a marketing channel is expensive, both at set-up and as an ongoing concern. One analysis estimates that transportation costs average 21% of the cost of goods sold in the United States; border-related trade barriers add a further 44%; and retail and wholesale distribution costs average 55% of the landed cost of goods sold in industrialized countries. These costs combine to a 170% total estimated trade cost above the cost of goods sold.[2] Clearly, managing the costs of such an effort can have a substantial impact on the product's ultimate sales and profitability in the market.

In addition, developing and running a marketing channel involves the use of other agents and/or companies (the designer's "channel partners") to whom the channel designer allocates important channel responsibilities, yet over whom the channel designer has limited control. This places the channel designer at considerable risk, because of the noncompensatory nature of the channel's work elements. In particular, it is typically not possible to compensate for a channel partner's nonperformance of a designated function by "doing more of" another function over which the channel designer does have control (imagine trying to compensate for a retailer's poor stocking efforts on store shelves by doing more promotion of the product, or shipping more frequently!). The successful channel designer and manager must therefore take account of this risk and foresee and manage any emerging problems; this requires a strategic and long-range focus rather than a tactical, short-term one.

Once established, marketing channel systems tend to be very difficult to dismantle or completely redesign. They involve investments in relationships with sales forces, distributors, and retailers; negotiations over shelf space or promotional considerations; and the development and adoption of information technology systems that allow the partners to communicate efficiently and quickly. These types of expenditures are not just current costs, but are rightly viewed as investments by the channel designer, with a payoff that is enjoyed for many years into the future (if the relationship is fruitful). The strategic goal is therefore to "do it right the first time."

Because of the positive externalities a well-run channel confers on other marketing efforts of the firm, its expense, the risk inherent in partnering with independent channel members, and the difficulty of redesigning or abandoning an established channel structure; it is therefore clear that proper channel design and management has *strategic* (i.e., long-run and wide-ranging) importance to the overall marketing strategy of the firm.

Demand-Side Insights for Marketing Channel Design

The fact that the end-user is the only channel member who *injects money into the channel system* enhances the value of understanding end-user demands and deciding how (and how intensively) to meet them. End-users demand not only products, but also a varied set of services from the channel. In short, it is not only of interest to a consumer *what* s/he buys, but also *how* s/he buys. The types of channel services demanded by consumers are sometimes aptly called "service outputs," identifying these services as the *output* of the work of channel members.

Service outputs vary across different specific channels and end-users. However, a generic list of service outputs includes *lot size*, *spatial convenience*, *assortment*, *speed of shopping/acquisition*, and various elements of *customer service*.[3]

Service outputs are inherently good. For example, avoiding shopping hassle, or getting a desired product quickly rather than with a delay, are attractive aspects of the buying process. However, not all consumers value a service output to the same extent. This implies that the delineation of consumer segments for marketing channel design is a function of the *variations* in consumers' service output demands. This is

quite a different concept of segmentation (*see* MARKET SEGMENTATION AND TARGETING) than the one a firm might use when thinking about physical product design; the fact that manufacturers send the same product to market through multiple channels implies that physical product attributes alone are insufficient to determine a product's sales and profit potential.

Thus, even a parity product can be significantly differentiated through the provision of carefully chosen service output levels. The strategic design of the service output bundle to offer along with the product can create *loyal* sales at a *high margin*, because the augmentation of the product with a desired set of service outputs effectively turns a competitive product into a highly differentiated product/service bundle. This is extremely good news for firms selling established lines of mature products that face head-to-head product-attribute competition. Indeed, it is a rare company that can boast more than one or a few truly innovative, unique, patent-protected products that face no current or potential competition. The ability to sell the usual product array depends on the channel's offering of an attractive combination of product and service outputs.

Supply-Side and Efficiency Insights for Marketing Channel Design

It might seem from the above discussion of service outputs that the channel should seek to provide all valued service outputs in order to prosper in a product-competitive marketplace. Unfortunately, this is not a generally feasible strategy, because of the cost of delivering intensive service output levels. Thus, while a demand-focused approach to marketing channel strategy is critical in developing consumer satisfaction, a complete *profit-maximizing* strategy requires controlling the cost side of the channel as well.

More generally, setting up a marketing channel structure involves addressing the following issues:

- decide what "work" must be done in the channel to generate targeted end-users' desired service outputs;
- choose channel intermediaries that can reach targeted end-users;
- split the "work" of the channel amongst the chosen channel partners for maximum efficiency (i.e., minimum cost);
- create a reward system that awards adequate compensation for channel members' work.

Decide what "work" must be done in the channel. The "work" of the channel includes performance of a set of *channel flows*, which are costly functions that must be done in a marketing channel in order to move the product (or service) from the hands of the manufacturer into those of the end-user. A generic list of channel flows includes *physical possession*, *ownership*, *promotion*, *negotiation*, *financing*, *risking*, *ordering*, and *payment*.[4]

Channel flows are *costly*; they can be done *by just one, or by more than one*, channel member; and they are done *in order to produce desired service outputs*. Because of the tension between the value of generating service outputs and the cost of doing so, it is worthwhile to weed out costly but non-value-added functions. One manager of a multinational company, sent to the South American marketplace with a mission of managing distribution costs, found that the company could remove inventory from over half of its stocking distributors in the multicountry region without compromising speed of delivery or product availability. An added benefit was the greater freshness of the product, which was subject to quality degradation over a period of months sitting in inventory. In such cases as this, careful monitoring and control of waste in the performance of channel flows can not only reduce costs, but can in some cases also actually *increase* quality and service provision.

In a specific channel application, the channel manager is urged to customize the generic list of channel flows so that their names and descriptions fit the market concerned. For example, *physical possession* and *ownership* may be fruitfully combined into one flow called *inventory holding* if the holder of product is commonly also its owner. On the other hand, it frequently makes sense to split up the *promotion* flow into multiple separate flows, such as *sales force costs*, *advertising*, *sales promotions*, and *trade show expenses*. Yet other flows, such as *risking*, may be renamed

to apply more directly to the business at hand; the cost of risking can be monetized by finding out the cost of an insurance policy to cover a negative outcome in the business, for example. If such is the case, *insurance* may be a better term than *risking* for that business.

Because distribution is so costly, it behooves the channel manager to try to audit the channel in order to understand channel cost levels, who bears them, and how to improve their management. It is often best to examine the accounting and financial systems of channel partners in order to see at what level channel costs can be measured.

When establishing a new channel structure or expanding an existing one, the channel designer must figure out what levels of channel functions need to be done in order to meet target end-users' demands for service outputs. For example, an arts and crafts retailer like Michaels in North America may first recognize that its "casual crafter" consumer segment wants quick availability of a wide range of arts and crafts products, along with project ideas and information on needed supplies and the appropriateness of the suggested project for various age and skill groups. These demands imply a high value for *quick delivery* (i.e., low *out-of-stock levels*), *assortment and variety*, and *customer service and education*. When Michaels wanted to grow its store network significantly in the 1990s and 2000s, the company realized that these service output demands translated into the need to manage the distribution channel from its suppliers, through its distribution centers, down to its retail stores, to improve the channel's performance on movement, storage, tracking, merchandising, and replenishment of product.[5] This recognition spurred a multiyear effort on the supply-side of the channel; such initiatives are substantial both in time and money costs. However, without such an effort, significant system growth would be impossible.

Choose channel intermediaries that can reach targeted end-users. Once the channel designer understands the work that needs to be done to generate desired service output levels, the structure of the channel must be determined. Channel structure specifies:

- the *types* of intermediaries to be used; and
- the *specific identity* of each intermediary with whom to partner.

There are many types of intermediaries from which to choose. For example, retailers may be bricks-and-mortar or online (or both); large and multistore or small and mom-and-pop in scope; organized as a franchise business; or broad-line or specialty stores. The manufacturer may use an independent distributor, a value-added reseller, an independent sales representative firm, or a sales broker to help sell its products to downstream buyers. Franchisees, direct selling individual resellers, or multilevel marketing distributor/resellers are alternatives to a standard retailer.

Similarly, once the manufacturer decides what types of intermediaries will be the best partners, it is also necessary to pick the specific partner(s) of each type. For example, when Jenny Craig weight-loss centers decided to augment its client service by offering to send weekly food shipments to them (rather than requiring them to stop in at Jenny Craig retail centers each week to pick up their food), it knew that it needed a third-party delivery partner (thus settling on the *type* of channel intermediary to use). Jenny Craig chose Federal Express as the shipper for its "Jenny Direct" program[6] – perhaps in a nod to FedEx's reputation for reliability and quick delivery, key service outputs for a consumer who is replacing all of their regular meals with Jenny Craig food.

More generally, in choosing both the *type* and the *specific identity* of its channel partners, a manufacturer generally takes into account the following factors:

- Market coverage (will the potential channel partner help me reach my target end-user segments?)
- Partnership potential and commonality of interest (will the potential channel partner benefit, as I will, from a strong effort to satisfy the target end-user segment?)
- Capability (can the potential channel partner perform the designated channel flows and functions using the right technology and with sufficient skill?).

Split the "work" of the channel amongst the chosen channel partners. With an identified set of channel partners, the channel designer must next decide what elements of the channel workload are to be allocated to which channel member(s). Some channel functions, such as inventory holding, may be performed serially by multiple channel members. Other channel functions, such as promotion or negotiation, could conceivably be focused in the hands of just one channel member. The goal in allocating the work of the channel is to control the total cost of running channel operations, because any unnecessary cost is lost profit to the channel as a whole. Another consideration (discussed further in the section below on channel conflict) is the safeguarding of the channel against the possibility that a channel partner shirks in its performance of channel workflow.

Create a reward system that causes channel members to earn adequate compensation for the work they do. A well-designed and implemented channel creates gross margins (revenue minus cost of goods sold) high enough to cover all channel costs, and to generate profit to each channel member. In order to preserve the incentives of each channel member to continue to do its allotted work and thus perpetuate good channel outcomes, adequate compensation must be awarded. The channel designer can use discounts off list price; functional discounts; spiffs; sales contests; and many other reward mechanisms to directly or indirectly influence the performance of appropriate channel effort.[7] A detailed analysis of incentive creation in the channel is beyond the scope of this discussion, but in general, good incentives are ones that:

- reward the desired behavior;
- can be recognized by the recipient as tied to the performance of the desired behavior;
- are of sufficiently high value to make it worthwhile to engage in the desired behavior.

Summary of supply-side and efficiency insights for marketing channel design. Note that in order to follow these supply-side suggestions for channel design, the channel designer ideally *knows* what work is in fact done by which channel members, and how much this work costs. In the real world, this can be a rather aggressive assumption, with neither of these dimensions generally well-known to all participants. However, it may still be possible to *estimate* the costs borne by channel partners, particularly in channel flows for which the channel captain bears a similar cost. For instance, a manufacturer may employ salespeople and thus have an idea of how much it costs its distributor to field a sales force; a channel partner's warehousing and trucking costs similarly may be estimated by a manufacturer's own costs in this area.

Ultimately, this suggests the value of developing metrics for the measurement of marketing channel activity and cost levels in order to manage channel costs, communicate the true cost of running the channel to stakeholders, and provide suitably motivational rewards to channel partners.

Marketing Channel Conflict as a Barrier to Effective Channel Management

Let us suppose that a well-designed channel structure has in fact been put in place, based on the analyses suggested above on the demand- and supply-sides. Alternatively, we can suppose that a preexisting channel has been analyzed and appropriate incremental changes have been made to close as many gaps on the demand- and supply-sides as possible. Can the channel designer now rest on his/her laurels as the channel seamlessly rolls along in the marketplace?

Unfortunately, the answer is probably no. Persistent *channel conflict* problems beset most channels, even very well-designed ones.[8] Channel conflict is characterized by one channel member's behavior which is in opposition to the desires of another channel member, so that the channel members are acting adversarially, rather than cooperatively, vis-à-vis one another. Three general types of channel conflicts are *goal*, *domain*, and *perceptual* (or *perception of reality*) conflict.

Goal conflict. Two channel members experience goal conflict when their objectives or time horizons are incompatible. This form of conflict

is ubiquitous in marketing channels, because of the very fact that channel partners are chosen for their *different* capabilities – not because they are clones of the channel designer. Although all channel partners may well be interested in the same higher-level goal, such as profit maximization, the actions that increase one channel member's profit do not always increase the profit of all channel members. For example, in a franchise channel, investments by the franchisee in better service, more personnel, training, cleanliness, etc. offer benefits to both the franchisee and the franchisor, because the resulting increased sales engender royalty payments to the franchisor as well as net income to the franchisee. However, the franchisee must bear all of these costs itself. Thus, the franchisee does not internalize the full benefits of these investments, and as a result, its goal of maximizing its bottom-line income (which includes those costs) will lead it to spend less on such investments than the franchisor would prefer.

Domain conflict. Domain conflicts refer to disagreements between channel members over either the *responsibilities of*, or the *rights from*, running the channel. Responsibilities include the performance of costly (but value-added) channel flows, and/or the adoption of appropriate technologies for getting the work of the channel done. Rights from running the channel include any direct or indirect financial benefits from participating in the channel. For example, a manufacturer selling B2B products to business customer end-users may use an independent distributor to reach its smaller customers, but reserve larger customers for its in-house sales force (*see* SALES FORCE STRATEGY). If the distributor invests in good faith in growing a customer relationship, and eventually succeeds in generating high sales from that customer, its "reward" may perversely be the removal of that customer from the distributor's control and reallocation of the customer account to an in-house salesperson. In this all-too-common example, the distributor exerts costly sales and support effort to build the customer account (i.e., takes *responsibility*), and as a result likely perceives a certain *right* to the profits from sales to that account. When the account is taken away, the distributor therefore experiences a domain conflict with the manufacturer, due to the distributor's belief that this customer was and remains part of the distributor's domain, while the manufacturer believes the customer rightfully belongs to the in-house salesperson.

Perceptual conflict. Perceptual conflicts between channel members occur when they view the same marketplace or set of market stimuli, but perceive them differently. Such conflicts are actually quite common; they can arise when a distributor or retailer has direct contact with end-users and therefore a deep understanding of what end-users want, while the manufacturer may not communicate with them directly. Another common source of perceptual conflict is a difference in norms or culture; what one country's manufacturer views as normal business practices (e.g., working late in the evenings or on weekends) may be viewed as unacceptable in some other countries or cultures. These perceptual differences can lead to conflicts over what constitutes adequate in-market sales effort or what promotional tools are the most effective.

Perceptual conflict is particularly insidious, because by its very nature, it is so hard to detect. The natural tendency is to assume that others see the world the same way that you do, so it can be very hard to recognize that a difference in perception is the root cause of conflict in the channel. If a conflict is not even recognized, there is little chance it will be resolved, much less, anticipated and prevented. Once perceptual conflict is recognized, it has often festered for so long that it is then difficult to deal with.

When any of these conflict sources is strong enough, the entire functioning of the channel may be in jeopardy, because different channel members specialize in the performance of certain subsets of channel functions and flows. A disgruntled channel member who shirks in the performance of its designated channel functions can therefore generate suboptimal service output provision and consumer satisfaction.

While some degree of conflict is to be expected even in a well-working channel, destructive conflict is to be avoided and, if possible, anticipated and prevented before it blossoms. For example, when John Deere began offering lawn tractors through Home Depot – even though it had historically always partnered

with its independent dealer network as its sole route to market – Deere management anticipated the possibility of serious domain conflict, because of the possibility of cannibalization (*see* CANNIBALISM) of dealer sales. To minimize this, Deere only offered entry-level products through Home Depot, while offering the full line through its dealers; as well, Deere dealers became the post-sale service provider for all Home Depot sales, thus offering a revenue stream even on units the dealer itself had not sold, as well as opening an opportunity to up-sell these consumers on accessories and, ultimately, their next lawn tractor. Dealers learned in this process that Home Depot sales did not just cannibalize their sales, but actually expanded the market, attracting new buyers who might never have visited a Deere dealership. Thus, even though adding another route to market is almost always a recipe for increased channel conflict, John Deere's actions prevented the conflict from escalating out of control. Indeed, the addition of the new channel in fact made the legacy channel members – Deere's independent dealers – actually *better off* than before the channel expansion.

In an example like this, the prevention of serious channel conflict involves not only an alert focus on channel partners' attitudes, but also an awareness of the tools available to the channel designer to combat emergent conflict situations. These tools are elements of channel influence or channel power, discussed in the next section.

Marketing Channel Power as a Tool of Conflict Management

When a channel member uses threats, monetary or nonmonetary rewards, contractual terms, education and expertise, or the value of the brand itself, to align the interests of its partners with its own, that channel member is acting as a channel captain and using sources of *channel power*. *Channel power* is defined as *the ability to induce another channel member to take an action which that channel member would not otherwise have done*.[9]

One conceptualization of power posits that channel member A's power over channel member B is greater, (a) the greater is the *utility* that A can offer to B, and (b) the *scarcer* is the source of A's utility to B.[10] The concept of scarcity is straightforward: the more A is like a monopolist, the more leverage A can exert over B. For instance, when establishing distribution in a foreign country, it is quite possible that there is only one distributor a manufacturer could use; this distributor wields considerable scarcity power. Meanwhile, the concept of utility simply means that A can offer B something that B values. This "something" could be market access, increased sales, technical expertise, or association with a valued brand-name partner. Further, these two fundamental power sources are complementary inputs to A's total power over B, so that each one enhances the incremental value of the other. A corollary of this is that without at least a modicum of one source of power, the value of the other power source is very low (e.g., with many rivals [i.e., low scarcity], a high level of utility does not afford A a high overall power position; analogously, if A's offer is not of high utility to B, it hardly matters if A is scarce or not).

While this notion of power arising from scarcity and utility is intuitively very appealing, it does not identify a specific typology of utility power sources. Another reference suggests five sources of power that can be thought of as elements of utility power: *coercive*, *reward*, *expertise*, *legitimate*, and *referent*.[11] Coercive power is the threat that a benefit will be withheld unless a channel partner behaves as the powerful channel member requests. Reward power, in some sense the converse of coercive power, is the offer of a valued benefit if the channel partner *does* behave as the powerful channel member requests. Expertise power is the ability to share valued knowledge with a channel partner; legitimate power is the perceived right that one channel partner has to influence another; and referent power is the ability to influence another channel member who wants to "identify with" the powerful channel member. By specifying *how* the powerful channel member offers utility to the dependent channel member, this framework identifies what type of influence one channel member has over others.

Importantly, these power sources describe the *potential* of a channel member to exert influence

over another; they do not require the actual exercise of that power. Also of importance is the fact that all channel members wield *some* power; by contrast, if a channel member had literally no power sources at all, it would exit the channel relationship because it could command no profit from the venture. Additionally, while one can audit a channel to evaluate current power positions of each channel member, a company can also make strategic investments in one or several power sources in order to develop its influence over time in its channel relationships. Examples of the accretion of power over time through such strategic investments are the accumulated brand equity (referent power) and operational expertise (expertise power) of franchisor McDonald's, or the reward power inherent in the size of a channel partner like Walmart.

In sum, channel power should be viewed as a *strategic resource* of the firm and investments in channel power should be made judiciously with a view toward the development of *scarce* sources of leverage.

Power can be abused as well as constructively used. In the short run, the use of channel power to improve one's own share of channel profits, to the detriment of other members' profits, may be successful. However, in the long run, this strategy can backfire for at least two reasons. One is that an overly greedy, coercive channel member may literally bankrupt other channel partners and cause them to leave the channel; this leaves the channel with some functions uncovered, which previously were done – essentially at subsidized levels – by the now-defunct channel partner. Ultimately these channel flows will have to be done by some other channel member, who is not likely to agree to subsidize their performance as the earlier channel member did. The other reason an abusive power-wielding strategy can backfire is that eventually, competitors will emerge with whom the "abused" channel partners will be more than happy to partner instead. Eventually, the abusive channel member will find itself without partners, both because its former partners desert it and because its reputation will precede it and make it difficult to form new relationships.

In sum, channel power is the overall term for a set of tools that give a channel member leverage over the actions of other channel members. In a constructive, well-working channel, power is either latent, or is exercised judiciously to further the channel's goals. In those situations of great power imbalance, the dominant channel member may take advantage of weaker ones in the short run, although in the longer run, such a channel structure is unstable.

Conclusion

Channel design and management is an important strategic focus of a successful market-focused company. It involves long-term attention to the demands of end-users not only for the firm's products, but also for the services the channel can offer; the necessity to manage the significant costs that distribution inflicts on the company, through judicious allocation of responsibility for channel flow performance to the chosen members of the channel; and the ability to maintain channel partners' positive incentives and motivations to execute on the specified channel design. These are formidable tasks for any channel designer or manager. However, the upside potential is great, precisely because the challenge is itself great; those who succeed are rewarded with long-standing and loyal relationships with both channel partners and end-user consumers and customers.

Endnotes

[1] Coughlan, Anne T. *et al.* (2006), p. 2; italics added.

[2] Anderson and van Wincoop (2004), p. 691–692. Their definition of "trade cost" includes "all costs incurred in getting a good to a final user other than the marginal cost of producing the good itself: transportation costs (both freight and time costs), policy barriers (tariffs and nontariff barriers), information costs, contract enforcement costs, costs associated with the use of different currencies, legal and regulatory costs, and local distribution costs (wholesale and retail)." The combined effect the authors calculate is $[(1.21 \times 1.44 \times 1.55) - 1] = 1.7$, or a 170 percent increase above production costs. Note that these estimates omit channel promotional costs, which would further increase the total cost estimate.

[3] See Chapter 2 of Coughlan *et al*. (2006) for a thorough discussion of service output demands.

[4] See Chapter 3 of Coughlan *et al*. (2006) for a detailed discussion of the eight generic channel flows.

[5] See Coughlan (2004) "Michaels Craft Stores: Integrated Channel Management and Vendor-Retailer Relations," Kellogg School of Management Case Series no. KEL036.

[6] See www.jennycraig.com for details on the "Jenny Direct" program.

[7] See Caldieraro and Coughlan (2007) for an analysis of the uses of spiffs in marketing channels; Caldieraro and Coughlan (2009) for a discussion of sales contests; and Chapter 11 of Nagle and Hogan (2006) for a discussion of some of the other reward structures mentioned here. Chapter 10 of Coughlan *et al*. (2006) provides legal insights into the use of various pricing and reward policies in a channel management context.

[8] See Chapter 7 of Coughlan *et al*. (2006) for a detailed discussion of sources of channel conflict.

[9] See Chapter 6 of Coughlan *et al*. (2006) for a detailed discussion of channel power.

[10] This notion of power is adapted to the channel context from Emerson (1962).

[11] French and Raven (1959).

Bibliography

Anderson, J.E. and van Wincoop, E. (2004) Trade costs. *Journal of Economic Literature*, **42** (3), 691–751.

Caldieraro, F. and Coughlan, A.T. (2007) Spiffed-up channels: the role of spiffs in hierarchical selling organizations. *Marketing Science*, **26** (1), 31–51.

Caldieraro, F. and Coughlan, A.T. (2009) Optimal sales force diversification and group incentive payments. *Marketing Science*, **28** (6), 1009–1026, DOI: 10.1287/mksc.1090.0493.

Coughlan, A.T. (2004) Michaels Craft Stores: Integrated Channel Management and Vendor-Retailer Relations, Kellogg School of Management Case Series no. KEL036.

Coughlan, A.T., Anderson, E., Stern, L.W., and El-Ansary, A.I. (2006) *Marketing Channels*, 7th edn, Pearson Education, Upper Saddle River.

Emerson, R.M. (1962) Power-dependence relations. *American Sociological Review*, **27** (1), 31–41.

French, J.R. Jr and Raven, B. (1959), The bases of social power', in *Studies in Social Power* (ed. D.Cartwright), University of Michigan, Ann Arbor, pp. 150–167.

Nagle, T.T. and Hogan, J.E. (2006) *The Strategy and Tactics of Pricing*, 4th edn, Pearson Education, Upper Saddle River.

marketing metrics

Donald R. Lehmann and Dominique M. Hanssens

Introduction

Marketing metrics are the measurements (typically quantitative) used to assess marketing effort and performance. These fall into three broad categories: capabilities, inputs, and outputs. Capabilities are the strength and resources possessed by an organization. Inputs refer to the effort and budget put into various marketing activities. Outputs refer to the impact of those activities at the customer, product-market, and financial-market levels.

Because marketing encompasses a broad spectrum of activities and capabilities, there is no such thing as "the" marketing metric. Rather, a variety of related measures are available to assess marketing capabilities, marketing inputs (activities), and marketing-sensitive outputs. Further, metrics have different purposes (e.g., descriptive vs. diagnostic vs. evaluative) and time frames (e.g., short- vs. long-run).

One convenient way to organize marketing metrics is in a five-component hierarchy (Figure 1). The first component is the set of capabilities available. The second component is what marketing does (i.e., marketing inputs). The other three components relate to the impact of marketing at the individual (customer), product-market (aggregate customer), and financial-market levels. The hierarchy implies that marketing activities are partially determined by capabilities and influence individuals who, in aggregate, determine product-market results (e.g., sales revenue or market share), which, in turn, drive financial results (including stock price).

Here we first discuss the different types of measurement that are used. We then describe the different levels of measurement. Next we present evidence of how the levels link together, with particular emphasis on the link of marketing activities to aggregate product-market and financial-market results. Here we use the term *products* to refer to either physical goods (e.g., an automobile) or services (e.g., an automobile insurance policy).

Figure 1 A simplified view of marketing metrics.

Types of Measurement

The term metric implies something can be measured, often quantitatively. Many types of measures are available, but most can be described on two dimensions: whether the measure is *direct* or *indirect* and whether it is *hard* (factual) or *judgmental*. For a concept such as sales revenue, a hard direct measure is available in company records, often as reported in annual reports, 10-Ks or 10-Qs. For more subjective concepts such as attitude (e.g., what customers think), a direct judgment measure is generally required. However, physiological measures may be utilized as well, including most recently fMRI measures designed to determine which regions of the brain have been activated.

In general, judgments involve *survey data*, such as customer ratings, hopefully based on sufficiently large and representative samples. Sometimes customer judgments are estimated or guessed by a firm insider (e.g., a manager or a researcher) or by an external research supplier. This practice is perilous, because individual human judgment is known to be subject to biases.

For some concepts, direct measurement may be either impossible to obtain or unavailable. In these cases, indirect measures are required. Indirect measures come in two types: formative and reflexive. *Formative* measures assess either

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

	Direct	Formative	Reflexive
Hard	Concept: sales Measure: units sold	Concept: customer lifetime value Measure: acquisition and retention rate	Concept: printer quality Measure: performance data on resolution, repair record, noise level
Internal judgment	Concept: market potential Measure: manager's assessment	Concept: salesforce quality Measure: salesperson performance ratings on timeliness, product knowledge, and customer friendliness	Concept: advertisement quality Measure: manager's assessment of ad likeability clarity, and presentation
External judgment	Concept: brand equity Measure: customer ratings of competing brands	Concept: competitive effort Measure: analysts' estimates of planned spending	Concept: software quality Measure: third party rating of software reliability, easy of use, and range of application

Figure 2 Types of metrics with examples.

(i) determinants or (ii) components of a concept. For example, quantitative and language skills are combined to "form" SAT scores to gauge a student's preparedness for college. Similarly, *reflexive* multiple questions about a brand (e.g., good–bad, necessary–unnecessary, desirable–undesirable) are used to "reflect" a consumer's attitude. By averaging the responses, a fairly accurate assessment of overall attitude toward the brand emerges. This assessment averages out slight differences in response to the separate questions as well as random error.

Overall, a direct, hard, and accurate (i.e., unbiased and with low random error included) metric is preferable. When one is not available, which other type of measure to use depends on a combination of availability, cost, and likely accuracy. Figure 2 provides a typology of the different types of measures available. Keeping these distinctions in mind, we now turn to a systematic discussion of the different levels of metrics in marketing.

Marketing Capabilities

What a firm does and how well it does it are largely a function of what (cap)abilities it has. The broadest capability is the firm's ability to relate to the outside world in general, and in particular to customers, that is, the firm's market orientation.

Market orientation has been defined in several different ways. Basically, it is the ability of a firm to (i) gather information on the outside environment in general and customers in particular; (ii) interpret, analyze, and understand the information, (iii) communicate the information throughout the organization; and (iv) make decisions based on the information and insights gleaned.

Other marketing capabilities relate to the ability of marketing to work with outside constituencies. This includes relations with

1. top management
2. other business functions such as R&D and sales (i.e., interfunctional coordination)
3. suppliers
4. channels
5. government

In addition, a number of capabilities relate to specific skills in the marketing organization, including the ability to

1. assess markets and research customer preferences
2. design and develop new products
3. price and market these products
4. advertise and promote
5. distribute
6. service

Some of these capabilities are readily assessed by simple quantitative measures, while others

are more qualitative and multidimensional in nature. As a result, we need to resort to different types of measurement to adequately represent the entire range of marketing capabilities. We discuss these different types first.

Input Measures

Input measures describe the effort put into marketing. To an important extent, marketing activities depend on marketing capabilities. These include specific abilities, for example, the ability to create new products, which is related to within-organization knowledge, experience, and access to external resources (e.g., through partnerships and collaborations). They also include broader organizational capabilities such as market orientation (generally considered to be the ability and effectiveness of a company in gathering information about and understanding customers, analyzing customer information and insights, and acting effectively on the analyzed information). Marketing capabilities develop over time as an organization "learns" through experience. While capabilities are critical inputs into decisions about marketing activities (also known as *programs*, *tactics*), here we focus on the activities themselves.

Product development. In addition to the overall R&D budget, which *types of products* and which *technologies or design approaches* (e.g., open source, collaborative development) are followed are key inputs.

Product quality. Products are often described by their *attributes*, which are relevant insofar as they provide tangible or intangible *customer benefits*. In some cases product quality is assessed externally, for example, by third-party comparative *product reviews*.

Service. *Service policies* (e.g., free returns, customer "loaners") are relevant inputs, as are *hiring and training policies* for customer-facing personnel.

Pricing. *Pricing schedules*, *discounts for quantity purchases*, and *price-matching guarantees* are key pricing decisions and marketing inputs. In the services sector, prices are increasingly expressed as "all-you-can-eat" *subscription* prices, notably in telecommunications and software. Also popular are *two* (or more) *part tariffs* which consist of a base fee and an incremental per use/product charge when use or purchases exceed a given threshold.

Channels. In addition to the overall budget for channels, *how it is allocated* to various forms of *channel support* (e.g., returns policies, cooperative advertising) is also a relevant input metric.

Personal selling. In addition to sales force budgets, which obviously impact the number of *sales calls made, compensation schemes* (e.g., salary vs. commission) and *hiring*, *training*, *and support programs* are critical inputs (although often controlled by a separate department within the firm).

Advertising and promotion. As for all activities, overall budget is a key input metric. *How the budget is allocated* is also critical. This is captured in the *media and promotional plans* (often represented in the form of "calendars"). Media include traditional customer touch points such as print, outdoor, direct mail, radio and TV, and the newer, interactive media such as Internet search, Internet display, and digital social networks. It also includes qualitative aspects such as *copy strategy*. Finally, some companies include *public relations* and *corporate sponsorship* efforts under advertising and promotion.

Output Measures: Activity Based

How current and potential customers (sometimes called *prospects*) respond to marketing activities is the first, and critical, indication of the effectiveness of marketing activities. Different activities have different initial impacts and hence measures of effectiveness. For example, for channels, *how many channels carry the product* (often assessed in packaged goods as percent ACV, the percent of all commodity volume sold by outlets that carry the firm's product(s)) is a basic measure. In addition, the (retail) *support they give the product* (e.g., featured in ads, displayed prominently, given multiple shelf facings, restocked regularly) is the result of both company efforts and customer demand. For a company with multiple stores (or franchisees), both the *number of*

stores and the *sales and margin per store* (also known as *same-store sales*) are critical metrics, since their product is total sales and margin respectively.

A more complete list of metrics reflecting the direct result of marketing activities appears in Table 1. Several of these results are of an *intermediate* nature, that is, they serve the purpose of assessing direct impact, but are not end goals in and of themselves. The link between intermediate metrics and overall business and financial performance is discussed later in the article.

Output Measures: Final Customers/Users

Because customers are critical to the market system, we discuss metrics related to them in more detail. Customer metrics can be described in a five-level hierarchy:

1. *Awareness*: The most basic level is awareness, that is, knowing that a product or brand exists and what it does or is used for. Both unaided (e.g., "which auto insurance providers have you heard of?") and aided (e.g., "which of the following auto insurance providers have you heard of?") *recall* are used to measure awareness. *Knowledge* about basic features is also a critical metric.
2. *Associations*: At least two categories of associations with a product (or brand) exist. The first is more factual and relates to *product performance* (e.g., the car handles well, gets good gas mileage, is easy to repair). The second is more psychological and *image based* and relates to *design* (e.g., the car is stylish), *users* (e.g., the car is for twenty-somethings), or *emotions* (e.g., the car confers status), as well as *personality* (e.g., the car is "macho").
3. *Attitude/Preference*: This level of metric assesses whether customers *like* the product, often measured on 0–100 scales or constant sum scales where points are divided among alternatives to reflect relative preference. Related measures are *intention to buy* and *willingness to pay* (WTP: the amount the customer is willing to pay for the product).
4. *Attachment*: *Loyalty* (or in the extreme, addiction) to a product is a key to repeat business, and hence future profits. The *extra amount one is willing to pay* for a particular alternative is one measure of loyalty, as is the *repeat purchase rate* (typically used in consumer goods businesses) and *the retention rate* (typically used in service businesses).
5. *Activity*: Customer *word of mouth* has long been an important manifestation of feelings toward a brand. This type of activity is also manifested in membership in product/brand groups (e.g., brand communities on the Web). Monitoring positive and negative comments is one way to keep track of this. A popular but controversial measure of activity is the *net-promoter score*, the number of people who would strongly and positively recommend a product or brand minus the number who would give a strong negative recommendation.

In addition to the metrics already discussed, four others are particularly important. While the first one is widely known (who has not been asked to fill out a satisfaction survey?), the others are also worth monitoring.

1. *Satisfaction*: Both *transaction satisfaction* (customer satisfaction with a particular transaction) and *cumulative satisfaction* are critical metrics. Event (transaction) satisfaction builds overall satisfaction, attitude, and intention to repurchase. Satisfaction depends on both the actual performance of a product and how well that performance compares to the customer's expectations of performance, the so-called "gap."
2. *Consideration*: A key metric is whether a customer is *willing to consider* purchasing a product or brand either actively (e.g., "what brands of tennis racquet would you consider purchasing?") or as a substitute (e.g., "if your favorite brand of paper towels – say, Bounty – were out of stock, would you consider buying Scott towels?" or "if Scott were on sale and Bounty was not, would you buy Scott?"). Interestingly, whether your brand is in a consideration set and how many other alternatives are in it as well gives an approximate prediction of market share as 1 (if your brand is in the set) divided by the total number of brands in the

Table 1 Specific marketing activity metrics.

Activity	*Input Measures*	*Output Measures*
Advertising	Media plan	Awareness
—	Copy	Attitude toward the ad
—	—	Attitude toward the product
—	—	Perceptions of the product
—	—	Reach (% who were exposed)
—	—	Frequency (average number of times people see the advertisement)
—	—	Cost per thousand (CPM)
Channels	Targeted outlets	% stores carrying
—	—	Retail support (features, displays)
Internet	Website design	Visits
—	Capacity	Pageviews
—	—	Click-throughs
—	—	Sales
—	—	Orders
Pricing	Base price	Price paid
—	Price schedule	Price premium
—	Price discounts	—
Promotions	Deal timing	Sales (percent) bought on deal
—	Depth of discount	Incremental sales
—	Method of distribution	Net profit: short run
Product development	Technology focus	Sales
—	Product focus	% sales from new products
Sales force	Time spent	Orders taken
—	Allocation of effort	Sales
—	Compensation method	Sales versus quota
—	Training	—
Service	Returns policy	Service satisfaction
—	Service center resources	Time to service
Overall strategy	Targeting	Sales
—	Positioning	Share
—	—	Profit/Margin

set. For example, if there are four brands in the set including yours, the probability a person will choose your brand is about 1/4.

3. *Confidence*: While attitude measures generally focus on a single estimate of how well a customer likes a product, how strongly they hold their opinions (i.e., their level of certainty) is also relevant, and correlates highly with actual choice.
4. *Brand value/equity*: Several commercial "mind-set" measures of the value of a brand to consumers are available. Most of their content is captured by the "5 As" (awareness, association, attitude, attachment, activity) discussed previously.

These final customer metrics are also intermediate output measures. For example, good service may enhance customer satisfaction, but this satisfaction in and of itself does not add value to the company. It needs to be translated to a metric of business or financial performance.

Product-Market Metrics

While metrics that assess customer attitudes provide primarily diagnostic information (i.e.,

why people do, or do not, buy a product), product-market measures effectively are evaluative (i.e., tell us overall how we are doing). Note that the metrics used to evaluate specific marketing activities include many such product-market-level measures.

Individual (customer)-level product-market metrics. For decades people in direct marketing have used their databases to assess their business performance with individual customers and prospects. Traditionally, they have followed an RFM approach: How *recently* a customer has made a purchase, how *frequently* they purchase (e.g., six times a year), and what the *monetary value* of their purchases is. The more recent and frequent the purchase and the greater the monetary value (preferably measured by margin rather than revenue), the more valuable is the customer. Effective marketing activities should increase at least one of these metrics.

More recently, the focus has shifted to assessing the value of a customer over his or her lifetime, also known as *customer lifetime value* (*CLV*). A current customer's future CLV is basically the total expected value of the net revenue (i.e., margin) received from a customer over his or her lifetime. In the special case, where the margin and retention rates are constant, a customer's CLV is $\mathrm{CLV} = \sum_{t=1} mr^t \left(\frac{1}{1+i}\right)^t = m\frac{r}{1+i-r}$. In this formula, r is the retention rate, that is, the probability that a customer in period $t-1$ remains a customer in period t, m is the marginal profit in a period from transactions with the customer, and i is the discount rate (typically 10–12%). Analyzing CLV at the individual customer level requires transaction records, something direct marketers (using mail, phone, internet) typically have available.

Overall (aggregate) product-market-level metrics. Overall performance is most typically measured in terms of *sales*, *share*, or *profits/margin*. Table 2 provides a list of common product-market metrics. The first group of aggregate metrics (sales, share, etc.) basically assesses sales and sales revenue (i.e., top-line measures), and are often contrasted with planned and/or expected results.

Table 2 Product-market-level output measures.

Individual-Customer Level:
Recency
Frequency
Monetary value
Retention
Margin
Customer lifetime value (CLV)
Aggregate Market Level:
Sales (unit, dollar)
Share
Share of wallet
Penetration/% buying
Price
Margin
Price sensitivity (elasticity)
Marketing sensitivity (elasticity)
Aggregate customer-based
Repeat rate
Retention rate
Margin per customer
Customer acquisition
Customer equity
Brand
Price premium
Volume premium
Revenue premium
New product introductions
Trial or adoption rate
Repeat rate
Volume
Growth pattern

The second group of metrics focus on the customer level. In aggregate, there are only three ways to increase revenue:

1. increase the retention rate of existing customers,
2. increase the margin (via up-selling or cross-selling) per customer, or
3. acquire new customers.

Both absolute values and relative values vis-a-vis competitors are worth monitoring. As an example, if two competing software providers A (the market leader) and B (a smaller player) are equally successful in acquiring new

customers, but B retains a higher fraction of its existing subscribers, then B will eventually overtake A.

If we take the sum of a firm's customers' individual CLVs, we obtain an important aggregate performance metric called *customer equity*; see Gupta and Lehmann (2005) for a detailed coverage. Like CLV, customer equity is a forward-looking metric that emphasizes long-term profitability as opposed to mere size. For example, a firm with a larger market share may have a smaller customer equity because its customers are not as profitable as those of a smaller competitor. Customer equity is related to a firm's stock price, as is discussed later in the article.

The final group of metrics assesses the relative performance of a brand. A brand with positive equity (i.e., one which is liked) can take advantage of this equity by virtue of a price premium (they can charge more than a generic competitor) or by virtue of extra sales (they have higher demand than the generic competitor). By combining these two sources of competitive advantage, we may compute the brand's *revenue premium* (the extra revenue a brand receives vs. an equivalent generic product), which is a useful measure of the value of a brand.

Assessing the Link between Marketing Activities and Product-Market Metrics

In order to make product-market metrics useful in marketing decision making, we must know how they respond to various marketing inputs or activities. Similarly, changes in intermediate performance metrics need to be converted to financial performance results in order to make them useful for resource allocation decisions. As different metrics can have widely different scales (for example, units, dollars, percentages, attitude ratings), we need a uniform linkage metric that enables comparison across different marketing inputs and outputs.

A classic (and very useful) measure of the effectiveness of an activity, or the business importance of an intermediate metric, is its *elasticity*. The elasticity of an activity is the percent change in a criterion variable Y (e.g., sales), which results from a given percent change in the activity X, written algebraically as: $\frac{\Delta Y/Y}{\Delta X/X} = \left(\frac{\Delta Y}{\Delta X}\right)\left(\frac{X}{Y}\right)$. Elasticities have the following desirable properties:

1. They are scale-free. For example, expressing marketing budgets in dollars, in thousands of dollars, or in Euros does not change their sales elasticities.
2. 'Percent change' language is easily understood by managers and researchers alike, and provides a useful input to marketing planning.
3. They represent three very different and strategically important response scenarios:

- diminishing returns, that is, elasticities between 0 and 1 in absolute value. When diminishing returns are present, marketing resources have to be allocated carefully, as overspending may result in substantial profit loss. Most marketing inputs (e.g., advertising) are subject to diminishing returns.
- increasing returns, that is, elasticities greater than 1 in absolute value. These often provide exceptional revenue generating opportunities, but are relatively rare.
- no returns, that is, elasticities are zero, an indication that the activity produces a deadweight loss.

While elasticities can be estimated subjectively, it is preferable to estimate them from relevant (e.g., past or comparable situation) data, for example via multiple regression analysis. This approach is sometimes referred to as *marketing mix modeling* or *market response modeling*; see Leeflang *et al.* (2000) or Hanssens, Parsons, and Schultz (2001) for detailed coverage. When X refers to price, you ideally want a small "up" elasticity (sales do not decrease much if you raise price) and a big "down" one (if you do cut price, you get a large increase in sales volume). That tends to be the case for strong brands, as opposed to their unbranded or weakly branded competitors. When X refers to a marketing budget item such as advertising or trade support, you ideally want the elasticity to be strong enough that it justifies the spending level. The most typical scenario is shown in Figure 3: the *sales response*

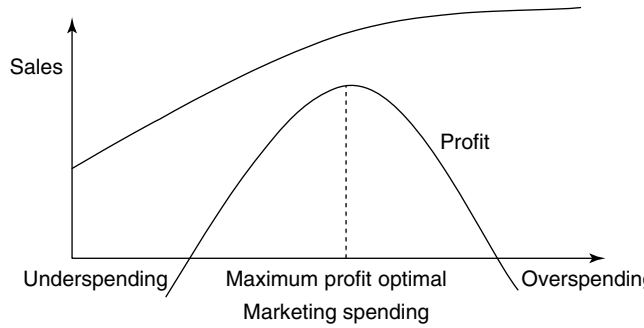


Figure 3 Typical sales and profit responses of marketing spending.

curve is concave, but the *profit response curve* is an inverted U. Knowing the shapes of these curves helps managers realize that higher revenues from more aggressive marketing do not automatically result in higher profits. The response curves also depend on the presence or absence of *competitive reaction*, and may be different in the short run versus the long run.

Interestingly, sales elasticities tend to be fairly consistent for the same marketing input across different business settings. Researchers have been able to derive "empirical generalizations" from many replications and these provide a starting point for subjective elasticity estimates, as well as a way to check the reasonableness of any given set of projections. We summarize below a few of these generalizations and refer to Hanssens (2009) for a more complete list.

1. *Advertising*: The short-term sales elasticity (impact on sales) of media advertising tends to be small (e.g., .01–0.03) for increases in spending for mature products that provide no new information about the product. The elasticities are more substantial (e.g., around 0.30) for ads for new products or for genuinely new uses of existing products. The long-run (total) effect of advertising has been found to be about 2 times the short-run effect for many consumer products.
2. *Personal selling*: The elasticity of sales calls is much higher than that of advertising, averaging upto 0.35. Like advertising, personal selling tends to be more impactful in the earlier stages of the product lifecycle.
3. *Pricing*: Estimating pricing elasticities is complicated by the fact that competition responds rapidly to price changes, which makes it difficult to disentangle the various influences. Nonetheless, price-level elasticities tend to be in the range of −1 to −3, with an average of −2.6 in competitive markets. Elasticities are much lower for monopolies or genuinely new products.

New-Product Performance

New products play a critical role in many businesses and early detection of likely success or failure is important. Beyond mind-set measures such as awareness and attitude, the first critical market level metric is *trial*. For frequently purchased products, *repeat* purchases are crucial for a product to succeed, (i.e., *adoption* occurs only for those who repeat buy a product) as is *volume per customer*. For durables (as well as for consumer products), the eventual number that adopt a product, that is, its *penetration level*, is crucial. In addition, the sales *growth rate* is important, because future earnings, once discounted to present value, become less relevant. Slow sales growth means future increases in sales will have less value.

Financial Metrics

Financial performance. Financial performance metrics (Table 3) are important if for no other reason than that accounting and finance professionals (e.g., the Chief Financial Offer or CFO) think in terms of them. Basic outcome measures include *revenue*, *net revenue*, and *cash*

Table 3 Financial metrics.

Performance
Revenue
Net revenue
Cash flow
Return on sales
Return on investment
Volume variance
Price variance
Valuation
Stock price
Stock return
Tobin's q

flow. Other measures assess the effectiveness with which resources (marketing budgets) are being deployed, that is, *return on spending* (*investment*). Pressure to demonstrate financial effectiveness is strong.

Other measures compare performance to planned levels, including the *volume variance*, that is, "actual versus planned unit sales," and the *price variance*, that is, the actual versus planned average price paid. Note these two are related because lowering price increases volume. Still there is some diagnostic value in learning whether a revenue shortfall (or overage) is due to reduced sales volume or a lower price. Albers (1998) expands these diagnostics to the other elements of the marketing mix.

Financial valuation. Linking marketing activities (e.g., new-product introductions) or assets (e.g., customer satisfaction, brand equity) to *stock price* is important but difficult. These links have been researched recently see Srinivasan and Hanssens (2009) for an extensive review. The basic logic underlying this work is the "efficient market hypothesis," which suggests that the stock market takes into account the future earnings performance of a firm and discounts it to present value to arrive at the price of a stock. (While few people believe all markets are instantaneously efficient in the near term, it seems reasonable to argue that stock markets do take into account future earnings.) As a result, (unexpected) changes in marketing actions can lead to unexpected changes in net revenue, which impacts stock value. While it is unreasonable to expect to detect the link between a minor tactical marketing activity and stock price (e.g., changing "vegetable soup/light" to "light vegetable soup"), there is evidence that more strategic marketing activities do impact stock price. Examples include new-product introductions and announcements, changes in advertising budgets, and customer satisfaction.

Note that for privately held firms or for nonprofits, stock price maximization is typically not the goal. For churches, for example, membership and attendance may be the key long-term performance metrics.

Summary

Many metrics are available to assess marketing performance, ranging from individual mind-set measures (awareness, associations, etc.) to changes in stock price. Measuring and striving to improve them can lead to improved business performance. Several caveats, however, apply:

Blind use of a few metrics can have unintended consequences: Use of a few key metrics can lead to efforts devoted to improving the metrics versus overall company performance. Further, it encourages "gaming the system," and runs the risk of encouraging manipulation of the metrics.

The distinction between forward- and backward-looking metrics is important: Most financial performance metrics are backward looking, that is, they describe what was, for example, sales or market share. More important for management are forward-looking metrics that are leading indicators of future performance, such as brand equity, customer satisfaction, and CLV.

Levels are interesting; changes and relative values are predictive: Knowing that, say, 80% of a firm's customers are satisfied is interesting. Knowing how this number compares to the competition, and whether it is increasing or decreasing, is important. More broadly, profitable revenue growth, in general, is a key objective. Among other things, it is reflected in a firm's price/earnings (P/E) ratio and stock price.

Means versus variances and customer heterogeneity: Both the level of variables and their variances matter. In particular, different segments of customers may have different views of a brand or a product, and these differences may be obscured in overall averages. Marketing metrics should be defined at appropriate segmentation levels, which, in some cases, means at the individual-customer level.

Bibliography

Albers, S. (1998) A framework for analysis of sources of profit contribution variance between actual and plan. *International Journal of Research in Marketing*, **15** (2), 109–122.

Ambler, T. (2003) *Marketing and the Bottom Line: The New Metrics of Corporate Wealth*, 2nd edn, Financial Times/Prentice Hall, Pearson Education, London.

Barwise, P. and Farley, J.U. (2003) Which marketing metrics are used and where? *Marketing Reports*, **2**, 105–107.

Farris, P.W., Bendle, N.T., Pfeifer, P.E., and Reibstein, D. (2006) *Marketing Metrics: 50+ Metrics Every Executive Should Master*, Wharton School Publishing, Upper Saddle River.

Gupta, S. and Lehmann, D.R. (2005) *Managing Customers as Investments: The Strategic Value of Customers in the Long Run*, Wharton School Publishing, Philadelphia.

Gupta, S., Lehmann, D.R., and Stuart, J.A. (2004) Valuing customers. *Journal of Marketing Research*, **41**, 7–18.

Hanssens, D.M. (ed.) (2009) *Empirical Generalizations about Marketing Impact*, Relevant Knowledge Series, Marketing Science Institute, Cambridge.

Hanssens, D.M., Parsons, L.J., and Schultz, R.L. (2001) *Market Response Models*, 2nd edn, Kluwer Academic Publishers, Boston.

Leeflang, P.S.H., Wittink, D.R., Wedel, M., and Naert, P.A. (2000) *Building Models for Marketing Decisions*, Kluwer Academic Publishers, Dordrecht/Boston.

Lehmann, D.R. and Reibstein, D.J. (2006) *Marketing Metrics and Financial Performance*, Marketing Science Institute, Cambridge.

Mela, C.F., Gupta, S., and Lehmann, D.R. (1997) The Long-term impact of promotion and advertising on consumer brand choice. *Journal of Marketing Research*, **34**, 248–261.

Moorman, C. and Lehmann, D.R. (2004) *Assessing Marketing Strategy Performance*, Marketing Science Institute, Cambridge.

Srinivasan, S. and Hanssens, D.M. (2009) Marketing and firm value: metrics, methods, findings and future directions. *Journal of Marketing Research*, **46** (3), 293–312.

Srivastava, R.K., Shervani, T.A., and Fahy, L. (1998) Based assets and shareholder value: a framework for analysis. *Journal of Marketing*, **62**, 2–18.

marketing mix

Gerald Albaum

The success of any MARKETING STRATEGY depends largely upon how the various marketing methods or tools are used by a firm to achieve a predetermined goal/objective. The planned and coordinated combination of such methods or tools to achieve goals and objectives is called the *marketing mix*, which is embodied in what is known as the firm's marketing program (Albaum and Duerr, 2008, p. 688). Thus, marketing strategy is reflected in how a firm integrates the elements of the marketing mix into the marketing program.

The marketing mix includes four major activities – product (*see* PRODUCT CATEGORY), price (*see* PRICING STRATEGY), distribution (*see* MULTICHANNEL MARKETING; MARKETING CHANNEL STRATEGY), and promotion (*see* INTEGRATED MARKETING COMMUNICATION STRATEGY; COMMUNICATIONS BUDGETING; SALES FORCE STRATEGY) – that a firm controls to meet the needs of customers (*see* CUSTOMER ANALYSIS; CUSTOMER SATISFACTION/DISSATISFACTION) within its target market (Pride and Ferrell, 2010, p. 5). A central feature is that it is built around the customer (or consumer). That is, the marketing program and the marketing mix should be formulated with the interests and needs of the consumer in mind. It must be structured in such a way that it integrates the customer into the company and leads to the creation and maintenance of a solid relationship between the company and the customer (*see* CUSTOMER RELATIONSHIP MANAGEMENT). A firm operating in this manner is said to be market driven or following the practice of MARKET ORIENTATION, and is concerned with what the consumer will buy that can be made profitably. The marketing mix is what makes all this happen. This is as true for international (including export) marketing (*see* GLOBAL MARKETING STRATEGY) programs and strategies as it is for domestic marketing programs and strategies.

In the last decade there have been a number of studies about various aspects of the concept of the marketing mix. A couple of representative examples are discussed below.

The choice of the elements of the marketing mix, or marketing tactics, is influenced by the firm's external environment. There are times when this environment can be complex and turbulent. As complexity increases, a manager's ability to plan, predict, and understand becomes harder, thus leading to more change. Turbulence involves fast and unexpected change in the environment, leading to greater uncertainty. What all this means is that in order for marketing success to occur in a turbulent environment, a different approach to the marketing mix is required. Mason and Staude (2009) propose a marketing mix tactics model based on chaos and complexity theories that can be used for success in a complex and turbulent environment.

Chen and Green (2009) studied the relationships between the marketing mix and customer-based PERCEPTION OF BRAND EQUITY between male and female shoppers in hypermarkets in Taiwan. Differences were observed between genders, with females scoring higher than males on customer-based brand equity. An earlier study of selected marketing mix elements and brand equity was reported by Yoo, Donthu and Lee (2000), who found how specific mix elements used can be either brand-building (high advertising spending) or brand harming (frequent use of price promotions).

Advances in analytical approaches to data analysis and the proliferation of media and channels of distribution have direct implications for the marketing mix and how it is managed in a company. Recent developments in modeling market response to marketing mix variables with specific emphasis on improving marketing resource allocation, given the above developments, is covered by Kerin and O'Regan (2008).

The concept of the marketing mix has stood the test of time since it was first popularized by McCarthy (1960) as the 4 P's! The mix elements are interrelated and interdependent, and can be substitutes and/or complements for each other. Industry has embraced the concept. For example, the trade publication *LP Gas Magazine* in early 2009 introduced a new section into its publication titled *Marketing Mix* (LP Gas, 2009, *p. 39*).

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Bibliography

Albaum, G. and Duerr, E. (2008) *International Marketing and Export Management*, 6th edn, Pearson Education Ltd, Edinburgh Gate, UK.

Chen, H.C. and Green, R.D. (2009) Marketing mix and branding: competitive hypermarket strategies. *International Journal of Management and Marketing Research*, **2** (1), 17–34.

Kerin, R. and O'Regan, R. (eds) (2008) *Marketing Mix Decisions: New Perspectives and Practices*, American Marketing Association, Chicago.

LP Gas (2009) *New Age, New Solutions*. March, 39.

Mason, R.B. and Staude, G. (2009) An exploration of marketing tactics for turbulent environments. *Industrial Management and Data Systems*, **109** (2), 173–190.

McCarthy, E.J. (1960) *Basic Marketing: A Managerial Approach*, Richard D. Irwin, Inc., Homewood, IL.

Pride, W.M. and Ferrell, O.C. (2010) *Marketing, 2010 Edition*, South-Western Cengage Learning, Mason.

Yoo, B., Donthu, N. and Lee, S. (2000) An examination of selected marketing mix elements and brand equity. *Journal of the Academy of Marketing Science*, **28** (2), 195–211.

marketing strategy models

Venkatesh Shankar

Introduction

Marketing strategy models are being increasingly developed and used by academics and practitioners worldwide. Be it a model for new-product forecasting or a model for allocating marketing resources, marketing strategy models form the basis for strategic marketing decisions.

A marketing model is a mathematical representation of a marketing phenomenon. A marketing strategy model is a marketing model focused on a strategic marketing problem (Shankar, 2008a). The main purpose of a marketing strategy model is to develop tools or insights for the formulation of marketing strategies.

For the purpose of this article, marketing strategy topics are viewed as those that exhibit the following characteristics: (i) the relevant marketing strategy decision variable has long-term consequences; (ii) the marketing strategy topic comes under the decision purview of senior marketing executives (e.g., chief marketing officer, vice president – marketing); and (iii) decisions relating to the marketing strategy topic involve a substantial amount of marketing resources (Shankar, 2008b).

Marketing strategy models can be classified into three types: descriptive models, normative models, and predictive models. A *descriptive model* is one that explains at least one variable (O) (e.g., sales, market share, abnormal financial returns) as a function of one or more marketing strategy variables (MV) (e.g., order of market entry, advertising spending) through a general equation such as $O = f(MV)$, where f is a linear or nonlinear function. A descriptive marketing strategy model may comprise one or more equations (system of equations). Typically, descriptive models are associative in nature; that is, a typical descriptive model simply states that two variables (e.g., a marketing strategy variable and sales) are correlated. Some descriptive models go beyond correlational relationships to explain a marketing phenomenon in terms of one or more marketing strategic variables causing the observed outcome. Descriptive models are typically calibrated (the parameters specifying relationships are estimated) using empirical data. An example of a descriptive model is a model of new-product spending by Gatignon, Weitz, and Bansal (1990). Advanced descriptive models include Bayesian models (see Rossi, Allenby, and McCullogh 2007).

A *normative marketing strategy model* is a model that offers prescriptions or guidelines to managers on what strategic decisions should be undertaken under specified conditions. For example, a normative model of market entry would provide managers with the recommended order and time of entry into a market for a given set of conditions. Typically, normative models are developed using optimization approaches from the operations research (OR) discipline or game theoretic principles. Usually, normative models are purely analytical in nature. That is, they are based on mathematical formulations, derivations, and expressions, and are not estimated using empirical data. However, some normative models are based on empirical data. A marketing allocation strategy model is an example of a normative model (Shankar, 1997).

A *predictive model* is one whose main purpose is to predict a future outcome. The future outcome could be strategically important to the firm and may or may not be a function of marketing strategy variables. A sales forecasting model is one such predictive marketing strategy model because sales forecasts are critical to the top-line performance of many companies. An example of a predictive model not based on a marketing strategy decision variable is the Bass (1969) model. This model predicts periodic sales as a function of cumulative sales. An example of a sales forecasting model that is based on marketing strategy variables is the channel sales forecasting model by Divakar, Ratchford, and Shankar (2005).

A typical marketing strategy model exhibits the following characteristics. (i) it contains important marketing outcome variable(s) and predictor marketing strategy variable(s); (ii) it comprises well-specified functional form(s) of the relationship between marketing strategy variable(s) and marketing outcome variables, based on a theoretical mechanism; and (iii) it is reasonably robust to different plausible real-world conditions.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

There are three basic types of empirical descriptive models depending on the nature of data used to estimate the model. If the data are cross sectional (e.g., from multiple firms at a single point in time), then cross-sectional models are most appropriate. If the data are longitudinal (e.g., from the same firm at different points in time), then time series models are most relevant (Dekimpe and Hanssens, 1999; Hamilton, 1994). If the data are both cross sectional and longitudinal (i.e., for a set of firms for multiple periods), then time series cross-sectional models are most commonly used. A special case of time series cross-sectional model is the panel data model, in which data come from the same set of cross-sectional units over time. Two types of panels exist. A balanced panel is one in which the time periods are the same for each cross-sectional unit. An unbalanced panel is a panel in which the time periods for some of the cross-sectional units are different.

Depending on the scale of the dependent variable, a marketing strategy model could be a regression model or a limited dependent variable (quantal) model. While most marketing strategy models are regression based, some models are limited dependent variable models, an example of which is a model of determinants of competitor response time by Bowman and Gatignon (1995).

Empirical industrial organization models are those that focus on marketing strategy decisions through analysis of both demand (e.g., sales, market share) and supply (e.g., pricing strategy, marketing allocation strategy) functions (Tirole, 1988). Typically, there is at least one demand equation and one supply equation. The demand and supply equations are derived from utility maximization principles and are hence structural in nature (Chintagunta *et al.*, 2006). In a marketing strategy context, the models are typically based on game theoretic analysis (Berry, Levinsohn, and Pakes, 1995; Shankar, 1997).

Consistent with the view of marketing strategy topics outlined earlier, marketing strategy decisions can be classified into three major topic areas: (i) new-product development (NPD) and product management strategy, (ii) market entry and growth strategy, and (iii) defensive or reaction strategy. (Table 1) Marketing strategy models relating to other topics such as customer equity (Reinartz and Kumar, 2003), brand equity (Shankar, Azar, and Fuller, 2008), and network effects (Shankar and Bayus, 2003), are important but are not discussed as they are outside the scope of this article. For detailed discussions on a more comprehensive set of marketing models, see Eliashberg and Lilien (1993), Hanssens, Parsons, and Schultz (2001), and Lilien, Kotler, and Moorthy (1992). Game theoretic models exist for analyzing competition in strategic marketing variables such as order and sequence of market entry, positioning, and defensive reactions, and in strategic levels of product, price, distribution channels, and advertising (for a detailed review, see Moorthy, 1985). For more details (*see* COMPETITIVE ANALYSIS).

The rest of the article is organized as follows. The basics of game theoretic models are summarized in the next section. Marketing strategy models under each topic are described in each of the remaining sections. Each section discusses the models used in representative studies, summarizes the findings from these studies, and outlines the strengths and weaknesses of the models.

Game Theoretic Models

Game theoretic models are being increasingly used for marketing strategy analysis. These models analyze strategic interactions between players in marketing strategy variables and predict market outcomes. In a marketing game theoretic model, the players (firms and/or consumers) play a noncooperative game in which they choose their marketing strategies. A firm's strategy is a complete specification of its actions under all possible contingent scenarios in the game framework.

The main assumptions of game theoretic models are that the players are rational and intelligent. In the context of firm marketing strategy, rationality refers to the firm's maximization of subjective expected utility (profits) and intelligence refers to the firm' awareness that its competitors are rational. In a game theoretic model, the goal is to identify the existence of equilibrium and solve for equilibrium strategies by the players. An equilibrium condition is represented by strategies from which

Table 1 Summary of selected marketing strategy models.

Paper	*Model Type*	*Data*	*Key Findings/Contribution*	*Key Limitations*
New-Product Development and Product Management Strategy				
Sorescu, Chandy, and Prabhu (2003)	Abnormal financial returns model	255 innovations from 66 firms during 1991–2000 from NDA pipeline	Dominant pharmaceutical firms introduce more radical innovations. Financial value of radical innovation increases significantly with marketing and technology support.	Nondominant firms may not have resources to spend simultaneously on multiple innovative ideas.
Grewal *et al.* (2008)	Regression model	308 firms in the Pharmaprojects database	Both early- and late-stage products-dominated portfolios are significantly positively associated with their financial values.	R&D intensity, advertising expenditures, and growth rate were omitted. Cross-sectional analysis had its drawbacks.
Market Entry and Growth Strategy				
Gatignon, Weitz, and Bansal (1990)	Regression model	68 brands of new drugs from 39 firms during 1978–82	Resources drive the level of launch communication efforts. Market growth rate moderates the effectiveness of detailing on market share.	The dynamics of competition or longitudinal data were not studied. The interaction of firm capabilities with environment was not considered.
Shankar, Carpenter, and Krishnamurthy (1998)	Trial, repeat generalized Bass model	13 brands from two categories from IMS Health database	Innovative late movers can enjoy advantage through greater market potential, higher repeat rate, and faster growth than the pioneer. They can also slow the growth and marketing effectiveness of the pioneer.	Only two product categories were analyzed.
Shankar, Carpenter, and Krishnamurthy (1999)	Dynamic sales response model	IMS Health data on 29 brands from six categories	Growth-state entrants reach asymptotic sales faster than pioneers or mature-stage entrants. Buyers are more responsive to pioneer's marketing spending. Mature-stage entrants are most disadvantaged.	The categories covered did not have failed brands.

Table 1 (*continued*).

Paper	*Model Type*	*Data*	*Key Findings/Contribution*	*Key Limitations*
Shankar (1999)	Rational expectations simultaneous equation model	23 new brands in six markets	New-product introduction spending is significantly influenced by incumbent reaction.	The product categories did not exhibit market exits.
Fischer, Shankar, and Clement (2005)	Relative market share model	73 brands in two categories from eight European markets	Late movers can reduce order of entry penalty by entering large countries in a sequential manner.	The US market was not included in the data.
Defensive Strategy				
Gatignon, Anderson, and Helsen (1989)	Regression model	Two new entries in the OTC-gyn category	Competitors retaliate (accommodate) with their most (least) effective marketing weapons.	Incumbent elasticities were not significantly altered by the new entrant. Uncertainty and the firm's perceptions were not included in analysis.
Shankar (1997)	New empirical industrial organization model	A large drug product category	Pioneers who adopt a follower (leader) role in a marketing mix variable in a static (growing) market and whose elasticity is decreased (increased) should accommodate (retaliate) against a new entry. Pioneers should accommodate (retaliate) with its competitively low (high) elastic variable.	Repositioning reactions were not considered. The response function was decoupled.
Shankar (1999)	Relative spending model	59 incumbent responses to entries in six categories	Higher spending by a new brand and multimarket contact result in weaker incumbent response.	Product and pricing reactions were not studied.

NDA, new drug application.

no competitor has an incentive to unilaterally deviate from its strategy.

Formulation and analysis of a game theoretic model involves many steps. The first step is to specify such rules of the game as the number of competitors, the set of feasible strategies or decision variable(s), the utility or profit function, the number of periods (single or multiperiod), whether one competitor is a leader or follower in the decision variable(s), and the information known to the competitors at different stages of the game. Some information such as which competitor moves first is known to all the competitors and is known as *common knowledge*.

If all information is common knowledge, then the game is known as a game with *complete information*. Some information such as manufacturing cost can be known to the focal firm but not to its competitors and are, therefore, labeled *private information*. In this case, some rules of the game are not common knowledge, so this game is known as a game with *incomplete information*.

Games can exhibit symmetric or asymmetric information. In the case of information symmetry, all competitors are similar in their information knowledge. In the case of information asymmetry, competitor X may have information on a variable (say cost) about competitor Y, but competitor Y may not have the information on the same variable about competitor X.

Different types of game theoretic models exist for analyzing different conditions. For example, a game in which one competitor is a leader and another is a follower is known as a *Stackelberg game* and the corresponding equilibrium is known as the *Stackelberg equilibrium*. In contrast, the equilibrium corresponding to a game in which all competitors move simultaneously is known as the *Nash equilibrium*. For the possible existence of both these equilibria, see Shankar (1997).

While much of game theoretic analysis is based on noncooperative interdependence among competitors, cooperative game theory examines the collusive behavior of firms. The predictions of game theory models can be tested through empirical analysis. While game theoretic models are powerful tools for competitive analysis, they suffer from limitations such as unrealistic assumptions about the game and the inability to capture real-world phenomenon through tractable models.

New-product Development and Product Management Strategy Models

Models on new-product development strategy include those on the cost and returns to spending on new-product development. Product management strategy models range from product selection models to product portfolio models. Although a number of models exist on this topic, only two are reviewed to offer a sample of relevant marketing strategy issues. Sorescu, Chandy, and Prabhu (2003) estimate a descriptive regression model of financial returns to radical innovations by pharmaceutical companies. They find that dominant firms introduce more radical innovations even after accounting for the source of the innovation (in-house vs acquired) and that they also introduce more market breakthrough and technological breakthrough innovations. Their results show that the financial value of a radical innovation significantly increases with greater product marketing, technology support, and product scope, and that it is significantly greater than the financial value of market breakthrough or technological breakthrough, but there is no significant difference between the values of market breakthrough and technological breakthrough innovations. Their model, however, does not address the risks associated with such innovations as well as with incremental innovations. Firms are interested in such issues as they decide on the number and mix of products in their NPD pipeline.

Grewal *et al.* (2008) relate the choice of products in the product portfolio by the stage of new-product development to the financial value of the portfolio. Their descriptive regression model has Tobin's Q as a function of net income, current ratio, debt-to-equity ratio, competitor's new-product portfolio market potential, the firm's new-product portfolio market potential, number of product categories, and cross-category variance. They estimate this model with data from the product portfolios of 308 firms in the *PharmaProjects* database.

They find empirical support for both early- and late-stage products-dominated portfolios to be significantly and positively associated with their financial values. Their model, however, does not include potentially relevant variables such as research and development (R&D) intensity, advertising expenditures, and growth rate due to nonavailability of data. Furthermore, the cross-sectional model that they use may not be able to fully capture the differences in effects as products move from one stage to the next stage over time.

Other models on this topic include those on new product alliances (Kalaignanam, Shankar, and Varadarajan, 2007), new-product announcements and preannouncements (Sorescu, Shankar

and Kushwaha, 2007), and product line strategy (Shankar, 2006).

Market Entry and Growth Strategy Models

Market entry and growth strategy models include those on decisions such as the order and timing of entry, forecasting, and new-product introduction. Market entry strategy involves decisions on the following questions: which markets should a brand enter and what growth strategy should it follow? When should a brand enter a particular market or country (order and timing of entry) and with what introduction strategy and degree of innovativeness? What are the predicted or forecasted sales for a new brand? In the international context, which countries should a new product or brand enter? How (in what order and with what marketing support) does or should the brand enter multiple countries?

To primarily address the effects of order of entry and innovativeness on a brand's sales, Shankar, Carpenter, and Krishnamurthi (1998) develop a brand level repeat purchase diffusion model with competitor diffusion effects. Their model is both descriptive and predictive. It posits brand sales as the sum of brand trials and repeat purchases. The trials follow a generalized Bass model and are a function of innovative and noninnovative competitor cumulative sales and total marketing expenditures of the various brands.

They estimate the model using data on 13 brands from two ethical drug categories. Their results show that innovative late-entrant brands can surmount pioneering advantage by enjoying a higher market potential and a higher repeat rate than the pioneer or noninnovative late movers, growing faster than the pioneer, slowing the growth of the pioneer, and reducing the pioneer's marketing mix effectiveness. The model used by Shankar, Carpenter, and Krishnamurthi (1998) and the results thereof provide insights into the order and timing of market entry and its trade-offs with innovativeness.

While all the previously discussed models capture the effects of innovativeness, order and timing of entry, and marketing efforts on sales, they do not explicitly address the issue of a brand's introduction strategy. Gatignon, Weitz, and Bansal (1990) specify a set of hypotheses about brand introduction strategies and identify the conditions under which a given marketing mix (excluding price) is more effective than another for continuous product innovations. They use a market share descriptive regression model in which brand market share is a linear function of relative product quality, detailing share, concentration ratio, number of brands (experience) in product category, and firm size. They also express detailing and the parameter associated with it as functions of exogenous variables such as market growth rate and market size. They estimate the model on cross-sectional data from 68 new drugs from 39 firms that were introduced during 1978–1982.

They find that the amount of communication efforts used to introduce a brand depends on the availability of financial resources. Their empirical analysis supports the importance of market growth and superior quality of the new product relative to existing products. It also shows that competitive structure of the market is extremely important, supporting the need for competitive analysis.

Their model and results were the first to shed light into an important phenomenon, but they did not analyze longitudinal data that are important in capturing the dynamic effects of launch strategies. They also do not capture the roles of order and stage of brand entry on sales into the launch strategy of a brand.

Shankar, Carpenter, and Krishnamurthi (1999) examine the effect of the stage of the product life cycle in which a brand enters on its sales through brand growth and market response, after controlling for the effects of order of entry and time in market. Their descriptive model of brand sales is a function of order of entry of the brand, time in market for the brand, cumulative sales of competitor brands, perceived product quality of the brand, and marketing mix expenditures of the brand and its competitors.

Estimating a dynamic brand sales model using 29 brands from six pharmaceutical markets, they report that the brand growth rate follows an inverted V pattern. They find that growth-stage entrants grow faster than pioneers and mature-stage entrants; competitor diffusion hurts the pioneer, has no effect on

growth-stage entrants, and helps mature-stage entrants; growth-stage entrants enjoy greater response to perceived product quality than pioneers and mature-stage entrants; pioneers enjoy higher advertising and sales force response than growth-stage entrants, which, in turn, have higher response than mature-stage entrants. They did not find a direct effect of order of entry or pioneering advantage.

These insights are useful for firms planning their entry and introduction strategies. Their data, however, did not include failed entries or exits from the market. Moreover, none of the previously discussed models consider the anticipated reactions of incumbents in deciding the launch strategy of a brand.

Shankar (1999) includes anticipated incumbent reactions in his model of the drivers of brand introduction strategy through an integrated framework that includes the determinants of both new-product introduction and incumbent response strategies. His joint model focuses on the interrelationship between new-product introduction and incumbent response strategies and on the role of multimarket contact in these strategies. Although the model is not normative, it has normative implications as it is based on rational expectations of firms. His new-product introduction model comprises both short-term and medium-term advertising and sales force expenditures of new brand as functions of the leadership of the brand in that marketing mix variable, size of the entrant, relative quality of the new brand, market experience of the entrant, anticipated reactions or marketing spending of large and small incumbent firms, multimarket contact of entering brand with incumbent firms, market size, and market growth rate. His incumbent reaction model consists of incumbent's ratio of spending in a marketing variable before and after the entry of the new brand, scale of entry of brand, brand dominance of incumbent at the time of entry of the new brand, and estimated elasticity of incumbent in the marketing variable after entry of the new brand.

He tested his model using US market data from several prescription drug categories. His findings show that new-product introduction strategy is influenced significantly by incumbent reaction strategy and vice versa. His results show that the relationship of a new product's marketing spending to the anticipated incumbent reaction differs across incumbents by size–anticipated reactions from large incumbents lead to low-entrant spending while anticipated reactions from small incumbents do not pose a threat to a new brand's spending. He finds that greater market experience helps reduce uncertainty about the effectiveness of marketing variables and contributes to greater entrant spending. By incorporating anticipated incumbent reactions, Shankar (1999) offers a model of new brand introduction strategy and important results. All the studies discussed thus far do not address decisions involving multiple international markets or countries.

With regard to international market entry decisions, there are two major decision variables: scope (extent of exposure across markets) and speed (how fast) of entry, leading to two possible strategies, namely, sprinkler (fast rollout) and waterfall (steady rollout) strategies. An example of a model of market entry strategy across international markets is a descriptive model of relative market share with endogenous entry and marketing mix variables (Fischer, Shankar, and Clement, 2005). This model captures the moderating effects of international market scope and speed of rollout of late-mover brands on their market shares relative to the pioneer in the focal country or market. They estimate the model by accounting for the endogeneity of international market entry strategy, order of entry, resources, quality, and other decision variables, as well as for unobserved effects, using pharmaceutical data on 73 brands from two product categories in eight European markets during 1987–1996. Their results show that broader international market scope is associated with a lower late-entry penalty and a greater marketing spending efficacy for late-mover brands. They find that speed of rollout, however, is unrelated to late-entry penalty, but a waterfall rollout strategy is associated with a greater marketing spending efficacy. They argue that late-mover brands that sequentially enter many large international markets can challenge the market pioneer in a country more effectively than other late-mover brands.

Taken together, the models offer some useful insights for marketing strategy. A brand's decision to enter a new market or country is based not only on the sales potential for that brand in

that market but also on the overall sales potential across multiple markets or countries and the regulatory processes in the different markets. A market pioneer grows faster, and enjoys greater repeat purchase rates than me-too late movers. Innovative late entrants, however, can grow faster than the pioneer and slow the pioneer's growth and marketing spending effectiveness. For a late mover, a strategy of entering the largest markets sequentially (waterfall strategy) may potentially reduce its late-entry penalty. Robust sales forecasting models can predict both trials and repeat purchases of new brands in the presence of competitive effects.

There are some areas that need further research and models as well. First, we need more models of market exits. Second, models capturing the effects of synergy across the brands of a firm on their market entry decisions merit future investigation. Third, more research is required on repeat purchase and brand diffusion models. Fourth, as product life cycle curves are different for different product categories, we need more empirical models across different product categories and different time periods to study these differences. Finally, more models on market evolution for new products and brands across countries are needed.

Defensive Strategy Models

Models on defensive strategy comprise competitor reaction or incumbent response models. Competitor response to new-product entry plays an important role in determining the evolution of the market. Both normative and descriptive models have been developed for response to new-product entry. Several normative or game theoretic models of incumbent response have been developed.

Descriptive models of competitor response to new entry have used econometric analyses of data from several industries. Gatignon, Anderson, and Helsen (1989) address how established competitors in an oligopoly react to a new-product entry in their market. They estimate an econometric model of demand response functions and reaction functions with data from the airline and over-the-counter gynecological (OTC-Gyn) product markets. Their descriptive model comprises brand market share as a function of advertising share and new market entries. They estimate the model with data from reactions to two new entries in an OTC-Gyn product market.

They argue that reaction time can be better understood and predicted by observing the effectiveness of a current competitor's marketing mix instruments. They find that incumbent firms react positively (retaliate) with their elastic marketing weapons and cut back (withdraw) with their inelastic marketing mix instruments. They assume that incumbent elasticities were not significantly altered by the new entrant.

Shankar (1997) develops a normative model of pioneer's defensive reaction and entrant's marketing mix spending. He develops a decoupled multiplicative sales response model in which a late-mover brand enters the market occupied by the pioneer and alters the pioneer's elasticity. The pioneer's sales response after the late mover's entry is specified as a function of advertising spending, sales force spending, price, and "time in market" of a new brand. By maximizing the profit function with respect to advertising, sales force spending, and price, he derives the equilibrium levels of spending for advertising and sales force. His result holds if both the pioneer and the late mover play a Nash game in all marketing instruments or if the late mover is a Stackelberg follower in one or all of the marketing instruments. He develops equilibrium reactions under Nash and different leader–follower games and empirically illustrates the analytical results with empirical analysis of data from a large pharmaceutical market.

On the basis of these results and the assumptions, he explains the pioneer's reactions and predicts its shift in marketing mix allocation upon new entry using empirical analysis of simultaneous and sequential games. He finds that the type of competitive game and the anticipated impact of the late mover on the pioneer's margin and elasticities are two critical factors that significantly affect the pioneer's decisions, in addition to the pioneer's characteristics and the market conditions considered by prior research. His results show that a follower (leader) role in a marketing mix variable, a static (growing) market, a decrease (increase) in own elasticity and margin generally lead to accommodation (retaliation) in that variable. He also highlights

cases in which general reactions do not hold and points out that it is necessary to look not only at one factor at a time but also examine the combination of all the factors. He argues that the shift in pioneer's equilibrium marketing mix allocation follows changes in its relative marketing mix effectiveness, which depends on the structure of competition, the impact of the late mover on its elasticities and margins, and the competitor's marketing mix.

Shankar's (1997) model and results offer interesting insights into incumbent reaction strategy. The defensive strategies proposed by his model are both theoretically and empirically driven, so they have important normative implications for managers. His empirical analysis, however, is based on one product category and his model does not include the role of multimarket contact in formulating the incumbent's defensive strategy.

Shankar's (1999) model, discussed earlier, incorporates the role of multimarket contact in incumbent's reactions. His results show that incumbent reaction is strongly related to new entrant spending. They show that incumbents react mildly to high entrant spending to avoid a competitive spending war and that multimarket contact between the incumbent and the entrant leads to milder incumbent response.

Collectively, these models offer some generalizable insights. When a new brand enters a market, incumbents retaliate with their competitively strong variables and accommodate with their competitively weak variables. An incumbent brand's defensive strategy is related to the entrant's new-product introduction strategy and vice versa. An incumbent tends to accommodate a new entrant when a new brand enters with a high level of spending and when they compete in multiple product markets.

There are many unexplored or underexplored issues relating to defensive marketing strategy models. First, models capturing the speed and duration of response will be useful for strategic decision making. Secondly, we need models of competitor reactions that capture multimarket competition across international markets. Thirdly, we require models to capture the reaction strategy of established brands when store brands or private labels are about to enter the market. Finally, we need more decision support systems (Wierenga, van Bruggen, and Staelin, 1999).

Conclusion

Marketing strategy models continue to grow in importance. We have gained a good understanding of some of important marketing strategy issues, such as, new-product development, product management, market entry and growth, and defensive strategies, through carefully developed and tested, descriptive, normative, and predictive models. The models range from standard econometric models to high-level game theoretic models. The models developed thus far have provided some generalizable insights into strategic marketing decisions. We need more models for international entry, allocation between R&D and marketing expenditures, product portfolio analysis, allocation across own marketing initiatives and strategic alliances and across own and comarketing spending, and greater decision support systems.

See also *competitive analysis; competitor analysis; marketing strategy*

Bibliography

Bass, F. (1969) A new product growth model for consumer durables. *Management Science*, **15**, 215–227.

Berry, S., Levinsohn, J., and Pakes, A. (1995) Automobile prices in equilibrium. *Econometrica*, **63** (4), 841–890.

Bowman, D. and Gatignon, H. (1995) Determinants of competitor response time to a new product introduction. *Journal of Marketing Research*, **32** (1), 42–53.

Chintagunta, P., Erdem, T., Rossi, P., and Wedel, M. (2006) Structural modeling in marketing: review and assessment. *Marketing Science*, **25** (6), 604–616.

Dekimpe, M. and Hanssens, D. (1999) Sustained spending and persistent response: a new look at long-term marketing profitability. *Journal of Marketing Research*, **36** (4), 397–412.

Divakar, S., Ratchford, B.T. and Shankar, V. (2005) *CHAN4CAST*: a multichannel multiregion forecasting model for consumer packaged goods. *Marketing Science*, **24** (3), 333–350.

Eliashberg, J. and Lilien, G. (1993) *Handbooks in Operations Research and Management Science: Volume 5, Marketing*, North-Holland, Amsterdam, Netherlands.

Fischer, M., Shankar, V., and Clement, M. (2005) *Can a Late Mover Use International Market Entry Strategy to Challenge the Pioneer?. MSI Report No. 05-004*, Marketing Science Institute, Cambridge, 25–48.

Gatignon, H., Anderson, E., and Helsen, K. (1989) Competitive reactions to market entries: explaining interfirm differences. *Journal of Marketing Research*, **26** (1), 44–55.

Gatignon, H., Weitz, B., and Bansal, P. (1990) Brand introduction strategies and competitive environments. *Journal of Marketing Research*, **27** (4), 390–401.

Grewal, R., Chakravarty, A., Ding, M., and Liechty, J. (2008) Counting chickens before the eggs hatch: on the valuation of new product development portfolios in the pharmaceutical sector. *International Journal of Research in Marketing*, **25** (4), 261–272.

Hamilton, J.D. (1994) *Time Series Analysis*, Princeton University Press, Princeton.

Hanssens, D., Parsons, L., and Schultz, R.L. (2001) *Market Response Models*, Kluwer Publishers, Boston.

Kalaignanam, K., Shankar, V., and Varadarajan, R. (2007) Asymmetric new product development alliances: win-win or win-lose partnerships? *Management Science*, **53** (3), 357–374.

Lilien, G., Kotler, P., and Moorthy, K.S. (1992) *Marketing Models*, Prentice Hall, Englewood Cliffs.

Moorthy, K.S. (1985) Using game theory to model competition. *Journal of Marketing Research*, **22** (3), 262–282.

Reinartz, W. and Kumar, V. (2003) The impact of customer relationship characteristics on profitable lifetime duration. *Journal of Marketing*, **67** (1), 77–99.

Rossi, P., Allenby, G., and McCulloch, R. (2007) *Bayesian Statistics and Marketing*, John Wiley & Sons, Chichester, England.

Shankar, V. (1997) Pioneers' marketing mix reactions to entry in different competitive games structures: theoretical analysis and empirical illustration. *Marketing Science*, **16** (4), 271–293.

Shankar, V. (1999) New product introduction and incumbent response strategies: their inter-relationship and the role of multimarket contact. *Journal of Marketing Research*, **36** (3), 327–344.

Shankar, V. (2006) Proactive and reactive product line strategies: asymmetries between market leaders and followers. *Management Science*, **52** (2), 276–292.

Shankar, V. (2008a) Strategic marketing decision models for the pharmaceutical industry, in *Handbook of Marketing Decision Models* (ed. B. Wierenga), Springer, New York.

Shankar, V. (2008b) Strategic allocation of marketing resources: methods and insights, in *Marketing Mix Resource Allocation and Planning: New Perspectives and Practices* (eds R. Kerin and R. O'Regan), American Marketing Association Publication, 154–183.

Shankar, V., Azar, P., and Fuller, M. (2008) BRAN*EQT: a model for estimating, tracking, and managing brand equity for multicategory brands. *Marketing Science*, **27** (4), 545–566.

Shankar, V. and Bayus, B.L. (2003) Network effects and competition: an empirical analysis of the video games industry. *Strategic Management Journal*, **24** (4), 375–394.

Shankar, V., Carpenter, G.S., and Krishnamurthi, L. (1998) Late mover advantage: how innovative late entrants outsell pioneers. *Journal of Marketing Research*, **35** (1), 54–70.

Shankar, V., Carpenter, G.S., and Krishnamurthi, L. (1999) The advantages of entering in the growth stage of the product life cycle: an empirical analysis. *Journal of Marketing Research*, **36** (2), 269–276.

Sorescu, A.B., Chandy, R.K., and Prabhu, J.C. (2003) Sources and financial consequences of radical innovation: insights from pharmaceuticals. *Journal of Marketing*, **67** (4), 82–102.

Sorescu, A.B., Shankar, V., and Kushwaha, T. (2007) New product preannouncements and shareholder value: don't make promises you can't keep. *Journal of Marketing Research*, **46** (3), 468–489.

Tirole, J. (1988) *The Theory of Industrial Organization*, The MIT Press, Cambridge.

Wierenga, B., van Bruggen, G.H., and Staelin, R. (1999) The success of marketing management support systems. *Marketing Science*, **18** (3), 196–207.

marketing planning

Nigel F. Piercy

The Process of Marketing Planning

Marketing planning is the organizational process consisting of the managerial activities, systems, analytical techniques, and formal procedures aimed at producing marketing plans. Marketing planning is a critical element of marketing strategy formulation and implementation; for example, see McDonald (2007). Planning involves a thorough strategic MARKETING AUDIT of the firm's strengths and weakness in relation to the opportunities and threats in a chosen target market (*see* SWOT ANALYSIS).

When marketing strategy extends beyond a single year, executives may develop a three- or five-year strategic marketing plan, as well as an annual plan to manage marketing activities during the year. Budgets for marketing activities are normally set annually. Planning frequency varies by company, business unit, market, and marketing activity. Market targeting and positioning choices are not changed significantly during the year, while tactical changes in product, distribution, price, and promotion strategies may be included in the annual plan.

The responsibility for marketing planning normally rests with a marketing executive or team. Regardless of format, the marketing plan is developed in close coordination with the strategic plan for the business. Increasing emphasis is being placed on involving all business functions in the marketing planning process, because coordination with other functions (R&D, finance, and operations) is essential (Hulbert, Capon and Piercy, 2003).

The Purpose and Scope of Marketing Planning

The marketing plan indicates the marketing objectives and strategy (*see* MARKETING STRATEGY), as well as the tactics for achieving objectives. Importantly, the plan provides a framework for *strategy implementation* (*see* MARKETING STRATEGY IMPLEMENTATION) and for control of marketing strategies and programs. Marketing plans are developed, implemented, evaluated and adjusted to keep the marketing strategy on target. Accordingly, marketing planning plays an integral role in the strategy process and is important to organizational learning capabilities (Piercy, 2009).

The plan represents the written statement of strategy and programs for each target market as a basis for actions. It may be the primary way for executives to obtain financial budget for a strategy, and to communicate the outcomes of strategy analysis to the organization.

In complex planning situations, the strategic business unit strategy indicates market target priorities, available resources, financial and other constraints, and other strategic guidelines required to develop marketing plans. The choice of planning unit will vary according to the company's product-market portfolio. Some firms plan and manage by individual products or brands, while others work with product lines, markets, or specific customers.

Content of a Marketing Plan

The format of marketing plans varies widely between organizations. Typically a marketing plan should specify the expected results (objectives), market targets, actions and responsibilities, schedules and dates, and control metrics. The plan should indicate details and deadlines, product plans, advertising and sales promotion actions, training requirements, and other information needed by line managers (for detail, see Cravens and Piercy, 2009).

Accordingly, an annual marketing plan should incorporate a *strategic situation summary* (describing markets, segments, competition and their strengths and weaknesses, and the company's competitive advantage in each segment of interest); *market target description* (defining and describing each market target and positioning strategy guidelines); *objectives for each market target* (objectives should be defined for each market target indicating financial performance, sales, market position, customer satisfaction, and other metrics); *marketing program positioning strategy* (how the firm wants to be *positioned* (*see* POSITIONING ANALYSIS AND STRATEGIES)) relative to the competition in the buyer's perspective, including statements of *product strategy* (setting the strategy for new products,

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

product improvements, and product deletions), *distribution strategy* (indicating the strategy for each distribution channel, including the role of channel members, assistance and support provided, and specific activities planned), *price strategy* (the role of price in the marketing strategy and actions planned) (*see* PRICING STRATEGY), and *promotion strategy* (indicating the planned strategy and actions for advertising, publicity, Internet, personal selling, and sales promotion); *marketing research* (information needs and planned projects); *coordination with other functions* (specifying responsibilities and activities of other departments that have an important influence on the planned marketing strategy); *performance metrics* (*see* MARKETING METRICS) (specifying the measures to be used to evaluate actual performance compared to the plan); and *forecasts and budgets* (predictions of revenues, profits, and costs to implement the marketing plan).

Managing the Marketing Planning Process

Executives' perceptions of the usefulness of marketing planning vary from favorable, to seeing planning as a bureaucratic ritual. These perceptions reflect how well the planning process is managed, and the link to strategy implementation.

A process perspective on planning emphasizes the need to design and manage planning process consistently to combine an analytical perspective (tools and techniques of analysis) with the behavioral dimension of planning process (executive perceptions, attitudes, participation, and assumptions) and the organizational context for planning.

Marketing Planning and the Strategy Implementation Process

Marketing planning should be evaluated in terms of achieving the outcomes specified in the plan. Nonetheless, the result of marketing planning is determined largely by the effectiveness of the strategy implementation process (*see* MARKETING STRATEGY IMPLEMENTATION). A good implementation process linked to marketing planning spells out the activities to be implemented, who is responsible for implementation, the time and location of implementation, and how implementation will be achieved. Two sets of factors are influential in determining implementation effectiveness: structural issues (including the company's marketing functions, control systems, and policy guidelines) and behavioral issues (concerning marketing executives' skills in bargaining and negotiation, resource allocation, and informal organizational networks) (Noble and Mokwa, 1999).

Bibliography

Cravens, D.W. and Piercy, N.F. (2009) *Strategic Marketing*, 9th edn, Irwin/McGraw-Hill, Burr Ridge, IL.

Hulbert, J.M., Capon, N. and Piercy, N.F. (2003) *Total Integrated Marketing: Breaking the Bounds of the Function*, The Free Press, New York.

McDonald, M. (2007) *Marketing Plans: How to Prepare Them, How to Use Them*, 6th edn, Butterworth-Heinemann, Oxford.

Noble, C.H. and Mokwa, M.P. (1999) Implementing marketing strategy: developing and testing a managerial theory. *Journal of Marketing*, **63** (4), 57–73.

Piercy, N.F. (2009) *Market-Led Strategic Change: Transforming the Process of Going to Market*, 4th edn, Betterworth-Heinemann, Oxford.

marketing strategy

Rajan Varadarajan

Introduction

Strategy exists at multiple levels in an organization, chief among them being strategy at the corporate, business unit, and functional levels such as marketing and R&D. Although the focus of this article is on marketing strategy, in light of the interdependencies between strategies at different levels in an organization, a brief overview of strategy at the corporate and business unit levels is presented first. This is followed by an in-depth discussion on marketing strategy focusing on the following issues: (i) perspectives on marketing strategy and the scope of marketing strategy; (ii) definition of marketing strategy; (iii) marketing strategy content and formulation process; and (iv) marketing strategy drivers and outcomes.

Corporate strategy refers to a firm's choice of businesses to be in. While some firms choose to remain as *single-business* firms, the prototypical, large firm of the twenty-first century in most regions of the world is a *multibusiness* firm. At the corporate level, a multibusiness firm's decisions regarding the choice of businesses to be in and the allocation of resources among businesses in its portfolio entail focusing on issues such as

- addition of new businesses to the firm's portfolio (diversification strategy);
- deletion of businesses from a firm's portfolio (divestitures strategy);
- mode of entry into new businesses (acquisition, internal development, joint venture, merger);
- mode of exit (sell-off, spin off);
- effective management of businesses retained (mission assignment and resource allocation among businesses in a firm's portfolio).

Business strategy refers to how a particular business in a firm's portfolio chooses to compete in the marketplace. The terms *business strategy*, *business unit strategy*, *competitive strategy*, *competitive business strategy*, *and market strategy* are used interchangeably in literature to refer to the strategy of a specific business unit in the portfolio of a multibusiness firm. The fundamental issue here is viewed as the manner in which the business achieves and sustains a competitive advantage (Teece, Pisano, and Shuen, 1997). A business achieves a competitive advantage by leveraging the distinctive skills and resources of the firm to implement a value-creating strategy that its competitors cannot implement as effectively (Barney, 1991). The skills and resources that a business can leverage to achieve a competitive advantage encompass those of the firm at large and are not limited to those that normally reside within the business (*see* COMPETITIVE ADVANTAGE: ITS SOURCES AND THE SEARCH FOR VALUE). Competitive advantage that is immune to erosion by competitive actions results in a sustainable competitive advantage. Viewed from the perspective of organizational functions, business strategy is multifunctional in scope.

Porter (1980, 1985) distinguishes between four broad generic strategies that entail different routes to achieving a sustainable competitive advantage as a basis for long-run, above-average performance:

- *Cost leadership*: achieving a competitive advantage through lower cost in a broad range of industry segments.
- *Differentiation*: achieving a competitive advantage through differentiation in a broad range of industry segments.
- *Focused cost leadership*: achieving a competitive advantage through lower cost in a narrow industry segment.
- *Focused differentiation*: achieving a competitive advantage through differentiation in a narrow industry segment.

Cost leadership strategy entails a business leveraging the distinctive skills and resources of the firm to achieve a defensible position of competitive cost advantage in the marketplace (i.e., being the lowest cost producer). *Differentiation strategy* entails a business leveraging the distinctive skills and resources of the firm to differentiate its product offerings from those of its competitors to achieve a defensible position of competitive differentiation advantage in the marketplace. Pursuing a cost leadership strategy does not necessarily mean selling the product at the lowest

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

price relative to competitors, but possessing the ability to compete on price, if warranted. In other words, cost leadership implies the inability of a business' rivals to compete on the basis of price in order to attract customers. However, this does not preclude a rival multibusiness firm that is able to achieve a competitive cost advantage from economies of scope from being able to compete on the basis of price. A differentiation strategy manifests as attributes or features in a business' product offerings that appeal to customers at large or to specific customer groups and for which they are willing to pay a higher price. In effect, the nonprice attributes of the business' product offering become more salient to buyers from the standpoint of brand choice decision. The effectiveness of a differentiation strategy is further enhanced when the market is segmented into homogeneous groups and distinctive value propositions are offered to each group (i.e., heterogeneity in customers' preferences is adequately met) (*see* MARKET SEGMENTATION AND TARGETING; POSITIONING ANALYSIS AND STRATEGIES; A FRAMEWORK FOR CREATING VALUE PROPOSITIONS).

There is considerable overlap in the domains of business strategy and marketing strategy, as conceptualized in management (strategic management) and marketing (strategic marketing) literature, respectively. Regardless, corporate, business, and marketing strategy interact together to shape the competitive positional (cost and/or differentiation) advantages of individual businesses in a firm's portfolio. Varadarajan and Clark (1994) provide an overview of the distinctive and overlapping domains of (i) corporate strategy and business strategy, (ii) corporate strategy and marketing strategy, (iii) business strategy and marketing strategy, and (iv) corporate strategy, business strategy, and marketing strategy. Varadarajan, Jayachandran, and White (2001) provide an overview of how a business' marketing strategy is influenced by and influences the firm's corporate strategy.

Perspectives on Marketing Strategy and the Scope of MarketinG Strategy

Broad to narrow: from comprehensive vector of decisions to a specific element of the vector. Marketing strategy, as an *umbrella* term, encompasses a number of strategy concepts, spanning from broad and comprehensive to narrow and focused marketing behaviors directed at consumers, customers, competitors, and/or other key constituencies in the marketplace. For example, the term *marketing strategy* is used in both marketing literature and marketing practice to refer to

- a comprehensive vector of marketing decisions encompassing numerous aspects of where to compete (e.g., markets to serve and market segments to target) and how to compete (e.g., multifaceted differentiation);
- a vector of marketing decisions encompassing numerous aspects of how to compete;
- a vector of marketing decisions concerning certain aspects of how to compete (e.g., push strategy versus pull strategy – pattern of allocation of resources among the advertising, personal selling, consumer sales promotion, and trade sales promotion elements of the promotion mix); and
- a marketing decision concerning a specific aspect of how to compete (e.g. market skimming price strategy versus market penetration price strategy, positioning strategy, and branding strategy).

Broad versus narrow: construal of marketing strategy as a single all-encompassing continuum versus one of two underlying continua. Defining marketing strategy necessitates addressing the question of the distinction between *marketing strategy* and *marketing tactics*. An examination of journal articles and marketing textbooks (textbooks on principles of marketing, marketing management, and marketing strategy) reveals diverse points of view regarding the purported distinction between marketing strategy and marketing tactics including the following: (i) the marketing behaviors of firms in the realm of the 4Ps (product, promotion, price, and place/distribution) are characterized as marketing strategy in some sources and as marketing tactics in other sources; in sources in the latter category, marketing behaviors pertaining to segmentation, target market

selection, and positioning (STP) are considered as the domain of marketing strategy and behaviors pertaining to the 4Ps as the domain of marketing tactics; (ii) in yet other sources, some elements of the 4Ps are characterized as pertaining to marketing strategy (product and place/distribution) and others as pertaining to marketing tactics (price and promotion); and (iii) in still other sources, certain marketing behaviors in the realm of each of the 4Ps are characterized as marketing strategy (e.g., promotion – push versus pull strategy; price – market skimming price strategy versus market penetration price strategy) and others as marketing tactics (e.g., promotion tactics and pricing tactics). Viewed against the backdrop of the dictionary *definition of definition* (an explanation or statement of the essential nature of anything), defining marketing strategy in terms of specific marketing behaviors that fall within its scope (e.g., STP-related behaviors of organizations) and those that fall outside of its scope (e.g., product, promotion, price, and place-related behavior of organizations) is problematic.

In the context of military and warfare, the concepts of strategy and tactics have an illustrious history. However, the adaptation of these concepts into some of the business disciplines including marketing has been haphazard and arbitrary. In reference to the delineation of the behavior of firms as strategic and tactical behaviors, Mintzberg 1987, p. 14) notes "*The point is that these sorts of distinctions can be arbitrary and misleading, that labels should not be used to imply that some issues are inevitably more important than others. . . . Thus there is good reason to drop the word "tactics" altogether and simply refer to issues as more or less "strategic," in other words, more or less "important" in some context, whether as intended before acting or as realized after it*."

The definition of marketing strategy presented here is based on the construal of the broad array of marketing behaviors of organizations as a *single underlying strategic continuum* rather than two distinct continua – strategic and tactical behaviors. All else being equal, it can be expected that the *resource commitments* associated with the marketing behaviors of organizations that are more strategic are (i) made with a relatively greater focus on facilitating the achievement of competitive advantage; (ii) relatively larger in magnitude; (iii) made with a relatively longer term orientation; (iv) made over a relatively longer time period, (v) relatively more difficult to reverse; and (vi) relatively more interdependent and internally consistent.

Broad and broadening: marketing strategy in an Internet-enabled market environment. Marketing strategy is generally viewed as principally concerned with an organization's decisions pertaining to where to compete and how to compete. Realistically, organizations in their own self-interest are likely to constantly explore and experiment with new ways of competing in the marketplace. Furthermore, developments in the macroenvironment often present organizations with new ways of competing in the marketplace, and thereby broaden the scope of marketing strategy (i.e., the vector of marketing decisions pertaining to how to compete in the marketplace). A case in point is the broadened scope of marketing strategy in the aftermath of the evolution of the competitive landscape from a predominantly physical marketplace to a marketplace encompassing both the physical and the electronic marketplaces (i.e., the Internet-enabled market environment), that is, a consequence of the competitive imperative to leverage the potential of the Internet to more effectively compete in the marketplace is the broadening of the scope of marketing strategy (*see* DIRECT AND INTERACTIVE MARKETING; E-COMMERCE AND INTERNET MARKETING). The change in the frame of reference for competitive marketing strategy from "How to compete in the physical marketplace?" to "How to compete in an Internet-enabled market environment that encompasses both the physical marketplace and the electronic marketplace?" has effectively broadened the vector of marketing decisions encompassing issues such as the following:

1. What should be the relative emphasis on traditional channels versus the electronic channel (i.e., the Internet) for
   - providing product-related *information* to customers;
   - *communicating* with customers;
   - *promoting* to customers;

- *transacting* with customers;
- *distributing* information products in digital form; and
- *customer trial/sampling* facilitation of information products in digital form.

2. What should be the relative emphasis on
   - marketing direct to customers via the Internet versus through intermediaries (traditional, electronic, and/or hybrid intermediaries) (*see* MULTICHANNEL MARKETING; DISINTERMEDIATION);
   - market pioneering (order of entry strategy); and
   - strategic alliances.
3. How can the potential of the Internet be leveraged to enhance innovations, customization, and augmentation in the product, price, promotion, and distribution elements of the marketing mix?
4. How can the potential of the Internet be leveraged to market an organization's offerings to new market segments and in new geographic markets?
5. How can the potential of the Internet be leveraged to market an organization's offerings in fundamentally new ways (new business models) (see Varadarajan and Yadav, 2002 for additional insights)?

Understandably, characterization of the prevailing market environment as an Internet-enabled market environment would be appropriate only in the context of markets in certain countries. Hence, the scope of marketing strategy (the breadth of the vector of competitive tools available to organizations to compete in the marketplace) is context dependent. Against this backdrop, a broad definition of marketing strategy and two somewhat focused definitions of marketing strategy are presented in the next section.

Definition of Marketing Strategy

The creation, communication, and delivery of products that offer value to customers in an exchange are central to various definitions of marketing. For instance, the American Marketing Association's (AMA) official definition of marketing reads as follows: "Marketing is the activity, set of institutions, and processes for creating, communicating, delivering, and exchanging offerings that have value for customers, clients, partners, and society at large (Marketing News, 2008, p. 28)." The creation, communication and delivery of products that offer value to customers in an exchange are also central to the definition of marketing strategy. At the most fundamental level, marketing strategy pertains to an organization's *decisions* that specify its *choice* of actions (from among alternatives) for creating, communicating, and delivering products that offer value to customers in an exchange.

At the broadest level, marketing strategy can be defined as an organization's integrated pattern of decisions that specifies its crucial choices concerning products, markets, marketing actions and marketing resources in the creation, communication and/or delivery of products that offer value to customers in exchanges with the organization and thereby enable the organization to achieve specific objectives. With reference to a specific product offering of an organization, the above definition can be stated somewhat succinctly as follows: *Marketing strategy refers to an organization's integrated pattern of decisions that specify its crucial choices concerning markets to serve and market segments to target, marketing actions to take, and the allocation of marketing resources among markets, market segments, and marketing actions toward the creation, communication, and/or delivery of a product that offers value to customers in exchanges with the organization and thereby enable the organization to achieve specific objectives* (see Varadarajan, 2010).

While organizations are faced with the need to address issues relating to "how to compete" on an ongoing basis and make midcourse changes as appropriate, the question of "where to compete" (choice markets to serve and market segments to target) is an issue that is addressed relatively infrequently. Hence, with specific reference to an existing product offering of an organization that is targeted at specific markets and market segments, the proposed definition can be stated even more succinctly as follows: *Marketing strategy refers to an organization's integrated pattern of decisions that specify its crucial choices concerning marketing actions to take, and the allocation of marketing resources among markets, market segments, and marketing actions toward*

the creation, communication, and/or delivery of a product that offers value to customers in exchanges with the organization and thereby enable the organization to achieve specific objectives (see Varadarajan, 2010).

The use of the terms *decisions* and *actions* in the definitions of marketing strategy merit brief elaboration. Within reason, the terms *actions*, *activities*, and *behaviors* can be used interchangeably. An organization's implementation of its marketing strategy *decisions* manifest as its marketing *actions* or marketing *activities* or marketing *behaviors* in the *marketplace*. Although a number of marketing *activities* also occur within the boundaries of an organization (e.g., new product development related activities; sales force training related activities), it is an organization's actions, activities, or behaviors in the marketplace directed at entities external to the organization such as its consumers, customers, and competitors that define its marketing strategy. In other words, it is to an organization's marketing actions, activities or behaviors in the marketplace that customers respond and competitors react (*see* Varadarajan, 2010 for a detailed discussion on the literature underpinnings of the above definitions and elaboration of the various constituent elements of the definitions).

Marketing Strategy: Content and Formulation Process and Drivers and Outcomes

Chief among the marketing strategy related questions of enduring interest to marketing educators and practitioners are the following:

1. What is marketing strategy (i.e., the *content* of marketing strategy, the primary focus of this article)?
2. What *processes* do organizations employ in order to make decisions concerning marketing strategy content (i.e., the marketing strategy formulation *process*) (*see* COMPETITIVE ADVANTAGE: ITS SOURCES AND THE SEARCH FOR VALUE; COMPETITIVE ANALYSIS; COMPETITOR ANALYSIS; MARKETING STRATEGY MODELS; MARKETING PLANNING; POSITIONING ANALYSIS AND STRATEGIES)?
3. What are the principal organizational and environmental *drivers* or *determinants* of an organization's marketing strategy (*see* MARKET ORIENTATION; MARKET/INDUSTRY STRUCTURE; SWOT ANALYSIS)?
4. What are some of the *outcomes* of marketing strategy that are of principal concern to organizations (*see* PERCEPTION OF BRAND EQUITY; CUSTOMER EQUITY; BRAND VALUE; MARKET-BASED ASSETS?

Table 1 provides additional insights into marketing strategy content and process-related issues as well as drivers and outcomes of marketing strategy. Specifically, the table highlights the focus of some of the other articles in this encyclopedia that relate to (i) marketing strategy content; (ii) marketing strategy formulation process, (iii) drivers of marketing strategy; and (iv) outcomes of marketing strategy.

Figure 1 presents a conceptual framework delineating the drivers and outcomes of marketing strategy (*see* Varadarajan and Yadav, 2002 (p. 298 to 309) for additional insights into the literature underpinnings of the links shown in Figure 1). As shown in the figure, chief among the factors influencing a business' competitive marketing strategy are the following:

1. distinctive skills and resources possessed by the firm (as noted earlier, the large, prototypical twenty-first century firm in most regions of the world is a multibusiness firm; hence, the term *firm* is used to refer to multibusiness firms and the term *focal business* or *business* is used to refer to a specific business in the portfolio of a multibusiness firm; in the case of single-business firms, the terms *firm* and *business* refer to the same entity);
2. structural characteristics of the market/industry in which the focal business competes;
3. characteristics of the focal business' product offerings;

Table 1 Marketing strategy: process and content, and drivers and outcomes.

Topic	*Focus: Marketing Strategy Content versus Process*	*Focus: Marketing Driver (D) versus Outcome (O) (see Figure 1)*
PERCEPTION OF BRAND EQUITY[a]	Content	D: Firm resource
CUSTOMER EQUITY[a]	Content	O: Outcome (Objective)
—	—	D: Firm resource
—	—	O: Outcome (Objective)
BRAND GROWTH STRATEGY	Content	—
BRAND STRATEGY	Content	—
BRAND VALUE	—	D: Firm resource
BUNDLING	Content	—
COMMUNICATIONS BUDGETING	Content	—
CANNIBALISM[b]	Content	O: Mkt. place & Fin. rerfor-mance
COMPETITIVE ADVANTAGE: ITS SOURCES AND THE SEARCH FOR VALUE	—	D: Firm resource
COMPETITIVE ANALYSIS	Process	—
COMPETITOR ANALYSIS	Process	—
CUSTOMER ANALYSIS	Process	—
THINKING DEEPER ABOUT CUSTOMER EXPERIENCE	Content	—
CUSTOMER LIFETIME VALUE (CLV)[c]	—	O: Financial performance
	—	D: Buyer characteristic
CUSTOMER RELATIONSHIP MANAGEMENT	Content	—
CUSTOMER SATISFACTION/DISSATISFACTION	—	O: Mkt. place performance
CUSTOMER SOLUTIONS	Content	—
DATABASE MINING AND MARKETING	Process and content	—
DEMAND ELASTICITY	Process	—
DIRECT AND INTERACTIVE MARKETING	Content	—
E-COMMERCE AND INTERNET MARKETING	Content	—
ETHICAL MARKETING AND MARKETING STRATEGY	Process and content	—
FIRST-MOVER (PIONEER) ADVANTAGE[d]	Content	O: Competitive advantage
GLOBAL MARKETING STRATEGY	Content	—
GO-TO-MARKET STRATEGY	Content	—
INNOVATION DIFFUSION	—	O: Mkt. place performance
INTEGRATED MARKETING COMMUNICATION STRATEGY	Content	—
LATER MOVER (NONPIONEER) ADVANTAGE[e]	Content	O: Competitive advantage
MARKET DEFINITION	Process	—
MARKET ORIENTATION[f]	Process and content	—
MARKET SEGMENTATION AND TARGETING	Process and content	—
MARKET SHARE	—	O: Mkt. place performance
MARKET/INDUSTRY STRUCTURE	—	D: Mkt./Ind. characteristics
MARKETING CHANNEL STRATEGY	Content	D: Market channel structure

Table 1 (*Continued*).

Topic	*Focus: Marketing Strategy Content versus Process*	*Focus: Marketing Driver (D) versus Outcome (O) (see Figure 1)*
MARKETING METRICS	—	O: Mkt. place performance
MARKETING MIX	Content	—
MARKETING STRATEGY MODELS	Process	—
MARKETING PLANNING	Process	—
MARKET-BASED ASSETS	—	D: Firm resource
	—	O: Mkt. Place & Fin. Performance
MASS CUSTOMIZATION STRATEGIES	Content	—
MULTICHANNEL MARKETING	Content	—
DISINTERMEDIATION	Content	—
POINT OF DIFFERENCE AND PRODUCT DIFFERENTIATION	Content	—
POSITIONING ANALYSIS AND STRATEGIES	Process and content	—
PRICING STRATEGY	Content	—
PRODUCT CATEGORY	—	D: Product characteristics
STAGES OF THE PRODUCT LIFE CYCLE	Process	—
PUSH AND PULL MARKETING STRATEGIES	Content	—
SALES FORCE STRATEGY	Content	—
SERVICES MARKETING STRATEGY	—	D: Product characteristics
SUPPLY CHAIN MANAGEMENT STRATEGY	Content	—
SWOT ANALYSIS	Process	—
TRADEMARKS, PROPRIETARY MARKS, AND BRANDS	—	D: Firm resource
A FRAMEWORK FOR CREATING VALUE PROPOSITIONS	Content	—
MARKETING WARFARE STRATEGIES	Content	

[a]Brand equity as a firm resource is a driver of marketing strategy (a resource that is leveraged to compete in the marketplace). Creating market-based assets such as brand equity, customer equity, and channel equity is an important objective of marketing strategy (i.e., an outcome in Figure 1). Leveraging the equity of a brand name in a firm's brand portfolio to enter a new product category relates to marketing strategy content. Similar reasoning is the basis for listing some of the other topics under multiple categories in the table 1.

[b]*Content*: Willingness to cannibalize as a deliberate strategy. *Outcome*: The financial and marketplace performance consequences of product cannibalization.

[c]Customer life time value is both a driver (*buyer characteristic*) and outcome (*marketplace and financial performance*) of marketing strategy.

[d]*Content*: First to market as a deliberate strategy. *Outcome*: Competitive positional advantages of market pioneering. Although competitive advantage is not shown explicitly in Figure 1 as an outcome of marketing strategy, chief among the objectives of marketing strategy is the facilitation of the achievement of competitive positional advantage (cost and/or differentiation advantage).

[e]*Content*: Early follower or late entry as a deliberate strategy. *Outcome*: Competitive positional advantages of later entry.

[f]*Process* and *Content*: Consider for instance Kohli and Jaworski's (1990, p.6) definition of market orientation: "Market orientation is the generation of organization-wide market *intelligence* pertaining to current and future customer needs, *dissemination* of the intelligence across departments and organization-wide *responsiveness* to it." The definition spans both process (intelligence generation and dissemination) and content (responsiveness). See Han, Kim and Srivastava (1998) for insights into the role of market orientation on the innovation strategy of firms.

4. characteristics of the buyers of the focal business' product offerings; and
5. macroenvironmental factors – economic, political and regulatory, social, technological, and so on – and changes in the macroenvironment (e.g., the impact of a technological development such as the Internet on how businesses compete in the marketplace).

The outcomes of marketing strategy are broadly delineated in Figure 1 as marketplace performance outcomes (e.g., market share, market share rank, customer satisfaction, customer loyalty) and financial performance (e.g., return on investment, customer lifetime value). Although not shown in the figure, it should be noted that marketplace and financial performance outcomes are also drivers of marketing strategy. Specifically, the impact of the marketplace performance and financial performance during time period, t, and earlier time periods on marketing strategy during time period, $t + 1$ are not shown in the figure. Also, in the interests of simplicity of exposition, mediated links between drivers of marketing strategy are not shown in the figure. Consider for instance, the following mediated link: macroenvironment characteristics → product characteristics → marketing strategy. The evolution of the market environment into an Internet-enabled market environment has had a major impact on the marketing strategy of producers of information products such as books, magazines, music, movies, newspapers, photographs, software, and videogames. Organizations extensively use the Internet as a distribution channel for information products in digital form as well as to facilitate product trial by consumers.

A business' marketing strategy decisions are influenced by both supply-side and demand-side considerations. In Figure 1, the links from firm characteristics, industry characteristics, and

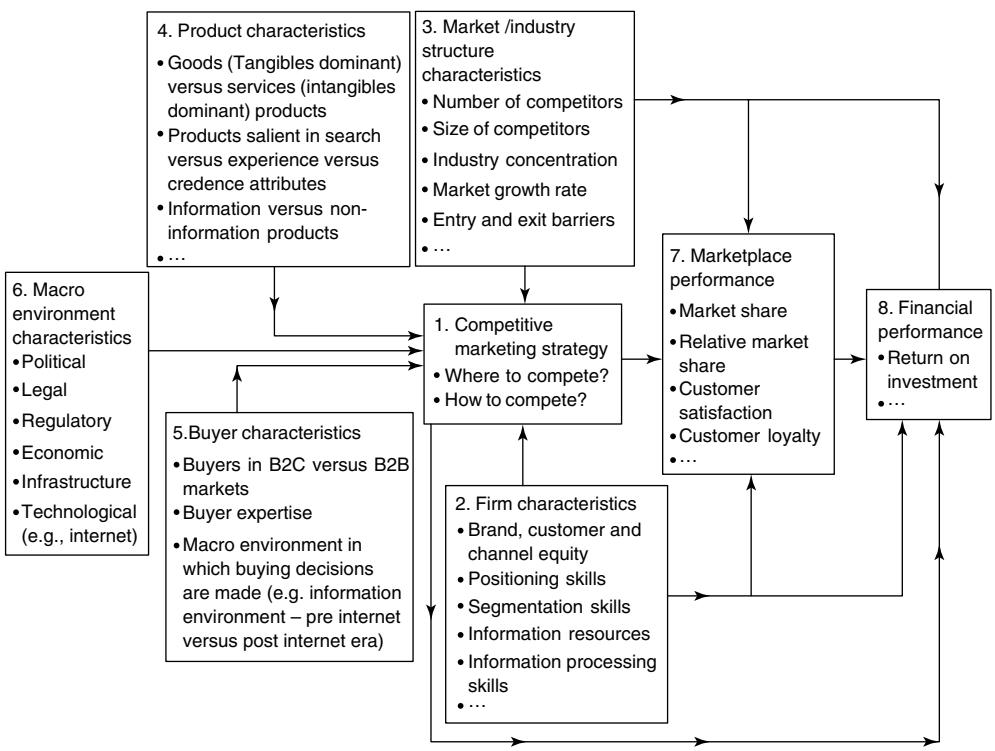


Figure 1 Marketing strategy: drivers and outcomes. (adapted from Varadarajan and Yadav (2002)).

product characteristics denote *supply-side drivers* of marketing strategy, and the link from buyer characteristics to marketing strategy denotes *demand-side drivers* of marketing strategy. The *structural characteristics of the industry* in which a business competes, a supply-side driver, refers to factors such as the number of competitors, size of competitors, entry and exit barriers, industry growth rate, and competitors' history of past behavior (*see* MARKET/INDUSTRY STRUCTURE). The *characteristics of a business' target customers*, a demand-side driver, refers to factors such as the attitudes, beliefs, and preferences, number and size, purchase frequency, sensitivity and responsiveness to various marketing instruments, and history of past behavior (*see* CUSTOMER ANALYSIS).

Of the supply-side and demand drivers shown in Figure 1, while the marketing strategy implications of firm characteristics is idiosyncratic to individual businesses competing in an industry (i.e., the marketing strategy implications of the heterogeneous resource endowments of competing businesses), the marketing strategy implications of the other three drivers are the same for all competing businesses. Consider, for instance, the growing importance of developing and nurturing two firm-specific skills and resources in the aftermath of the evolution of the market into an Internet-enabled market environment:

- *Information resources*: the nature and amount of information acquired and accumulated by the organization about individual customers.
- *Information technology resources and information processing skills*: the ability of the organization to use information technology (IT) resources and information processing skills to gain insights about individual customers and use this knowledge to customize its future interactions with these customers.

In an Internet-enabled market environment, an organization's information resources, information technology resources, and information processing skills are key *strategy enablers* in that they determine its ability (or lack thereof) to pursue certain marketing strategies as well as make more informed decisions in the realm of various marketing-mix variables. For instance, an organization's ability to engage in *suggestive selling* (recommend products to an individual customer based on perceived similarity with the purchase patterns of other customers) depends on its information resources, IT resources, and information processing skills. Similarly, in the area of pricing, an organization's resources and skills in the above areas determine the extent to which it would be able to engage in more fine-tuned *price discrimination* (e.g., inferring buyers' price sensitivity from their web navigation and purchase behavior and offering products at prices customized to the level of individual buyers). An organization's information resources, IT resources, and information processing skills are also a critical determinant of its ability to pursue other strategies in the realm of pricing such as *dynamic pricing* (e.g., changing the price at which a product is offered on the basis of prevailing supply and demand conditions) (*see* Varadarajan and Yadav, 2002 for additional insights) (*see* DATABASE MINING AND MARKETING).

Conclusion

Marketing strategy refers to the *behavior* of an organization in the marketplace. Broadly construed, such behavior is *multifaceted* (e.g., positioning, differentiation, pricing, advertising, and distribution related behaviors), *coordinated, integrated*, and *internally consistent across multiple facets* and *directed at entities external to the firm* (e.g., end-use customers, intermediate customers, consumers, and competitors) with a clearly defined *purpose* (i.e., to achieve specific organizational objectives). The bases for an organization's behavior in the marketplace are its *decisions* pertaining to *where to compete* and *how to compete*. The implementation of an organization's marketing strategy *decisions* manifest as its marketing *actions*, *activities*, or *behaviors* in the marketplace.

Bibliography

Barney, J.B. (1991) Firm resources and sustained competitive advantage. *Journal of Management*, **17** (1), 99–120.

Han, J.K., Kim, N., and Srivastava, R.K. (1998) Market orientation and organizational performance: is innovation a missing link? *Journal of Marketing*, **62**, 30–45.

Kohli, A.K. and Jaworski, B.J. (1990) Market orientation: the construct, research, propositions, and managerial implications. *Journal of Marketing*, **54**, 1–18.

Marketing News (2008) Marketing defined. *Marketing News*, **42**, 28–29.

Mintzberg, H. (1987) The strategy concept I: five Ps for strategy. *California Management Review*, **30**, 11–24.

Porter, M.A. (1980) *Competitive Strategy*, Free Press, New York.

Porter, M.A. (1985) *Competitive Advantage*. Free Press, New York.

Teece, D.J., Pisano, G., and Shuen, A. (1997) Dynamic capabilities and strategic management. *Strategic Management Journal*, **18** (7), 509–533.

Varadarajan, R. (2010) Strategic marketing and marketing strategy: conceptual domain, definition, fundamental issues and foundational premises. *Journal of the Academy of Marketing Science*, forthcoming.

Varadarajan, R. and Clark, T. (1994) Delineating the scope of corporate, business and marketing strategy. *Journal of Business Research*, **31**, 93–105.

Varadarajan, R., Jayachandran, S., and White, J.C. (2001) Strategic interdependence in organizations: deconglomeration and marketing strategy. *Journal of Marketing*, **65**, 15–28.

Varadarajan, R. and Yadav, M.S. (2002) Marketing strategy and the internet: an organizing framework. *Journal of the Academy of Marketing Science*, **30**, 296–312.

See also *a framework for creating value propositions; brand growth strategy; brand strategy; brand value; bundling; cannibalism; communications budgeting; competitive advantage: its sources and the search for value; competitive analysis; competitor analysis; customer analysis; customer equity; customer lifetime value (CLV); customer relationship management; customer satisfaction/dissatisfaction; customer solutions; database mining and marketing; demand elasticity; direct and interactive marketing; disintermediation; e-commerce and Internet marketing; ethical marketing and marketing strategy; first-mover (pioneer) advantage; global marketing strategy; go-to-market strategy; innovation diffusion; integrated marketing communication strategy; later mover (nonpioneer) advantage; market definition; market orientation; market segmentation and targeting; market share; market-based assets; market/industry structure; marketing metrics; marketing mix; marketing planning; marketing strategy models; marketing warfare strategies; marketing warfare strategies; marketing channel strategy; mass customization strategies; multichannel marketing; point of difference and product differentiation; positioning analysis and strategies; pricing strategy; product category; push and pull marketing strategies; sales force strategy; services marketing strategy; stages of the product life cycle; supply chain management strategy; SWOT analysis; thinking deeper about customer experience; trademarks, proprietary marks, and brands*

market-based assets

Tasadduq Shervani

Market-based assets arise from the commingling of the firm with entities in its external environment. Brands, customer relationships, distribution channel, and other partner relationships are all examples of market-based assets (Srivastava, Shervani, and Fahey, 1998). Each such asset requires a focal firm, an entity from its external environment, and a relationship between the two that creates value for the focal firm.

Market-based assets are becoming more critical for firms as their value has become more dependent in the last 50 years on intangible assets, or assets that are not captured on the balance sheet of a firm. In fact, for many firms, market value is a multiple of their book value. It is critical for such firms to understand and explain the gap between their market value and book value. Although there are other types of intangible assets that can account for the gap between market and book value (e.g., intellectual assets such as patents), market-based assets play an important role in explaining this gap. In contrast to intellectual assets, market-based assets are relational assets that exist because of key relationships between a firm and its customers and partners.

These relational assets create value for a firm because they allow a firm to (i) accelerate cash flow, (ii) enhance cash flow, (iii) reduce the volatility and vulnerability of cash flows, and (iv) increase the residual value of cash flows. For example, brands can lead to faster trial and adoption of new products introduced by a firm, thus accelerating cash flows. Strong customer relationships can increase customer loyalty and retention, which can reduce the vulnerability of cash flows. A well-established and well-nurtured customer base is a socially complex, difficult-to-imitate, and relatively rare asset that can increase the long-term residual value of a firm.

The ultimate goal of the marketing function within a firm can be conceptualized as increasing the value of the market-based assets of the firm. Therefore, marketers must first assess the current value of their firm's market-based assets. The next step is to leverage market-based assets to create shareholder value by accelerating, enhancing, protecting and stabilizing cash flows. Finally, the effect of marketing actions to leverage such assets must be measured. This, in turn, will give marketers permission to invest in the next round of marketing programs and activities to strengthen their market-based assets and increase shareholder value.

Market-based assets have existed for as long as business has existed. However, the tendency in the past was to lump them together with all other intangible assets under the category of "goodwill." Goodwill was seen as a reward for having run a business successfully in the past. By understanding that market-based assets are a significant portion of the goodwill earned by a firm over time and by understanding the factors that lead to the development and exploitation of market-based assets, marketing researchers and practitioners can better understand how marketing creates value.

The relationship between marketing and finance, the two critical functions in any business, can be better managed using the common language of market-based assets. Finance personnel can now understand the value created by marketing in a language that they are familiar with. Marketers, in turn, can use this common language to ensure that the firm makes an appropriate investment in the marketing function and marketing activities.

Bibliography

Srivastava, R.K., Shervani T.A., and Fahey F. (1998) Market-based assets and shareholder value: a framework for analysis. *Journal of Marketing*, **62**, 2–18.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

mass customization strategies

Arun Sharma

The concept of mass customization is becoming increasingly important to businesses. To illustrate, on May 20, 2009, Amazon.com listed 46 books with the name mass customization in their title and 3000 books with mass customization as key words. Google Scholar listed more than 2000 articles and books with mass customization in their title. Mass customization is the strategy by which firms provide customers with products that are individually tailored to their specific needs. Mass customization strategies are extensively used in two domains, manufacturing and marketing. In manufacturing, the strategy refers to production techniques that allow firms to customize products at efficiencies of mass production. In marketing, mass customization is the strategy used in the practice of customer-centric marketing. Sheth, Sisodia and Sharma (2000) stated that in customer-centric marketing, marketers first assess each customer or potential customer individually and determine whether it is more profitable to serve that customer directly or through a third party. On the basis of that determination, customer-centric marketers then decide whether to create a customized or standardized product offering and/or some other element(s) of the marketing mix.

Mass customization requires modular and rapid assembly processes and therefore the shift from mass manufacturing to mass customization requires redesigning products and production processes. Consumers are generally aware that major firms such as Dell, Nike ID, and Land's End practice mass customization. Other firms that are less commonly known to practice mass customization include McDonald's and Chinese restaurants. Chinese restaurants are a classic example of mass customization strategies; a limited number of ingredients allow a Chinese restaurant to create an unlimited number of meal variations in a very rapid time frame.

Mass customization marketing strategies emerge when firms change from a supply-side orientation to a demand-side orientation. Mass customization marketing strategies also emerge when customers' need for customization is high and matches marketers' ability to modularize production (Figure 1). In cases where marketers can customize but customers have minimal customization needs or desires, segmentation marketing strategies emerge (e.g., shampoo and detergent markets). When customers need or desire customization but marketers do not have the ability to modularize, component customization marketing strategies emerge. Examples of these marketing strategies are found in the clothing (alterations are done at the store) and automotive (dealer-installed accessories) industries. Finally, when marketers' ability to customize is low and customers do not need or want customization, mass marketing strategies dominate.

Even within mass customization, different marketing strategies emerge (Figure 2), based on whether customization is driven by a marketer or by customers. If the extent of customization is high by both a marketer and its customers, a "cocreation" marketing strategy often emerges. In cocreation, a marketer and its customers jointly develop a product offering. Cocreation marketing strategies have traditionally been limited to business-to-business marketing, but recently cocreation marketing strategies have been applied in consumer markets such as computers and automobiles. When customers drive customization and marketers do not, customizable products emerge. Examples of these types of products are software (customized by users) and cellular phones (customized ring tones, etc.). In the case of high marketer-driven customization and minimal customer-driven customization, these marketing strategies are characterized by environmentally driven products. Examples are gasoline (temperature-based additives), food (customized to region), and clothing (regional fashions).

In summary, mass customization is a strategy that enables the practice of customer-centric marketing. To successfully apply a mass customization strategy, marketers need to shift from a supply-side orientation to a demand-side orientation and change their product development and manufacturing processes.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

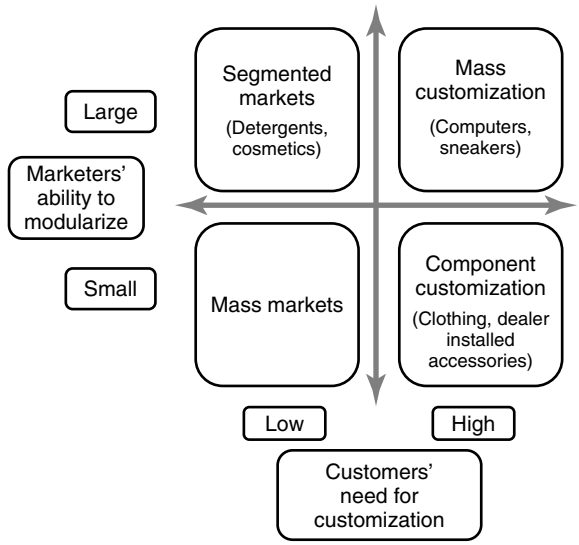


Figure 1 Types of markets.

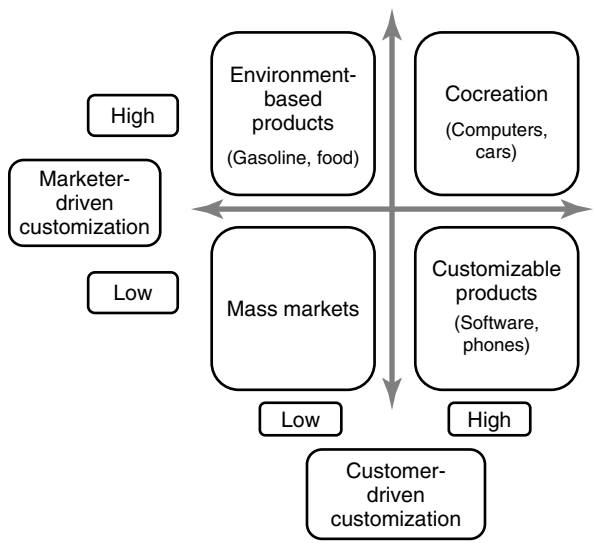


Figure 2 Types of mass customization.

Bibliography

Sheth, J.N., Sisodia, R. and Sharma, A. (2000) The antecedents and consequences of customer-centric marketing. *Journal of the Academy of Marketing Science*, **28** (1), 55–66.

multichannel marketing

Adel I. El-Ansary

Multichannel marketing is marketing through multiple channels to reach target market segments. Marketing channels may be classified in terms of formats initiated by manufacturers, wholesalers, retailers; or for that matter other channel members who perform marketing functions, participate in marketing flows, or facilitate marketing transactions. Retail-based channel formats are of particular interest to marketers because of their constant and direct connection to end user customers in general and consumers in particular. Retail formats consists of store retailing and nonstore retailing such as web sites (*see* MARKETING FUNCTIONS ON THE INTERNET), mail order catalogs, direct mail, and newer forms of technology such as text messaging and email. Traditionally, marketers refrained from the use of secondary channels, in order to avoid explicit conflict with primary channel members and implicit confusion in projecting brand image to their customers. Today, the emergence and continued growth of "multichannel customers," led marketers to view multichannel marketing as the norm, rather than the exception (Coughlan *et al.*, 2006). When customers use different channels to search for information, compare brands and prices, and conclude related transactions, marketers find themselves compelled to use any combination of channels necessary to effectively communicate with their customers and render the purchase process more efficient.

Multichannel marketing is a powerful tool for businesses. Channels can interact and reinforce each other, ultimately giving marketers greater opportunities to network with existing and potential customer bases. Fostering business-customer interaction (*see* CONSUMER INVOLVEMENT) is one of the main purposes of multichannel marketing. Such interaction allows for greater promotion of products and services, leading to increased profits for the company, greater customer loyalty, reduced costs, and enhanced branding (*see* BRAND EXTENSIONS AND FLANKER BRANDS). Multichannel marketing also enables marketers to track customer behavior, thus creating a better understanding of the customers' *decision* (*see* CONSUMER DECISION MAKING) processes. By understanding the decision process, businesses can improve customer retention rates (Rangaswamy and VanBruggen, 2005).

There are two facets of multichannel marketing: the supply side and the demand side. Suppliers find multichannel marketing attractive because of increased market penetration (*see* SUPPLY CHAIN MANAGEMENT STRATEGY) and the potential to reduce competition through barriers to entry. Suppliers have access to various markets and segments thereof through multiple channels. The end consumer also gains from multichannel marketing. Multiple channels empower consumers to select channels that meet their demand (*see* INTEGRATED MARKETING COMMUNICATION STRATEGY) for service level outputs such as assortment, lot size, and delivery time. Also, by pitting channels against each other, consumers are able to demand more services at lower prices. Perhaps, the greatest contribution of multiple channel marketing is the ease of discovery. Suppliers and customers alike are able to locate each other with ease, and meet each other's demands effectively (Coughlan *et al.*, 2006).

There are many factors that a marketing manager must consider when determining the appropriate channels to reach customers. Managers must determine who the potential customers are (*see* CONSUMER CATEGORIZATION), what products are those customers purchasing, when are they purchasing products, where are they buying products, and how are the products being purchased. Multichannel marketing is viewed differently by diverse firms. Some firms consider multichannel marketing as a means of diverting lower-valued customers to low-cost channels, such as web sites, in order to keep transaction costs down. However, other firms view multichannel marketing as means of reaching higher-valued customers (*see* SOCIAL CLASS).

There are challenges that arise when a firm decides to implement multichannel marketing. Aside from the potential of conflict, intrabrand competition and channel *cannibalization* (*see* CANNIBALISM) are problems that result in lost sales to competitors. Moreover, inconsistent *customer experience* (*see* THINKING DEEPER ABOUT CUSTOMER EXPERIENCE)

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

with various channels can negatively impact overall brand image (Eisenberg, 2009).

Bibliography

Arikan, A. (2008) *Multichannel Marketing: Metrics and Methods for on and Offline Success*, Sybex.

Coughlan, A.T., Anderson, E, Stern, L.W., and El-Ansary, A.I. (2006) *Marketing Channels*, 7th edn, Pearson Education.

Eisenberg, B. (2009) Multichannel marketing challenges: part 1. www.ClickZ.com, May 2.

Rangaswamy, A. and VanBruggen, G.H. (2005) Opportunities and challenges in multichannel marketing: an introduction to the special issue. *Journal of Interactive Marketing*, **19**, 2.

Sauer, A., El-Ansary, A.I., Torsten Tomczak, T. and Reineke, B. (2007) Managing Change in Distribution, in *Multichannel Marketing Handbook* (ed. B.W. Wirtz), Gabler-Springer Science, pp. 221–234.

Wilson, H., Street, R., and Bruce, L. (2008) *The Multichannel Challenge*, Butterworth-Heinemann.

disintermediation

Rajendra Sisodia

Disintermediation refers to the removal of existing intermediaries from a supply chain (*see* SUPPLY CHAIN MANAGEMENT STRATEGY). Some refer to disintermediation simply as "cutting out the middleman." The World Wide Web has made it cost effective for many companies to deal with more customers directly and has contributed to the trend toward disintermediation.

Traditionally, marketing has depended heavily on the presence of multiple intermediaries between a producer and consumer. These intermediaries add time and place utility to the functional utility "engineered" into the product or service offering of a producer. In addition to serving as physical conduits for getting products to the market, intermediaries also serve as informational conduits, conveying relevant information from producers to customers and customers to producers.

The advent of the World Wide Web changed the role of intermediaries in many industries. The Web enabled a high degree of market transparency, leading customers in some industries to initiate the process of disintermediation in search of lower prices. Simultaneously, a company's virtual presence on the Web enabled a high level of accessibility almost immediately, and also enabled two-way information flow directly with end-users (customers). Companies could serve huge numbers of customers efficiently and effectively by automating many administrative tasks by means of the Web. This led to rapid growth in support services, such as customer relationship management.

A well-known example of *downstream* disintermediation is Dell, Inc., which achieved dramatic growth in the 1990s by selling direct to consumers and businesses. Wal-Mart has been successful at *upstream* disintermediation; by dealing directly with producers, it lowered its product costs and passed the savings along to customers.

Despite these well-known examples, the extent of Web-driven disintermediation has been less than was initially expected. This is because customers continue to value many of the functions provided by retailers and wholesalers (such as credit, the ability to experience products, a wide assortment of offerings, and returns processing). Companies have also found that it is often less efficient to ship individual orders than shipping products in bulk to stores. Many retailers have also sought to enhance their own value to customers by integrating their physical and virtual presence (referred to as *bricks and clicks*). This has endeared them to customers who like being able to buy online and pick up or return products at a physical store, or vice versa.

Reintermediation

Reintermediation refers to the emergence of new types of intermediaries between end-users and producers after disintermediation has already occurred. New categories of intermediaries arise to capture value-creating opportunities through direct interactions between customers and producers. As with traditional intermediaries, these new types of intermediaries thrive based on the "economic transfer" principle: those that deliver greater value at lower cost tend to prosper.

For example, Priceline has emerged as a new type of market intermediary, using a "reverse auction" method to bring buyers and sellers together. The company refers to its patented business model as *buyer-driven commerce*. Potential buyers submit bids known as *conditional purchase offers* to buy products such as airline tickets at a certain price; sellers can either accept the offer, reject the offer, or counteroffer. Other examples of new types of intermediaries include rating services, automated ordering services, and services based on consolidating numerous small orders from many customers into more economically viable quantities.

In summary, the phenomena of disintermediation and reintermediation have both been driven by the dynamic and growing capabilities of the World Wide Web. The default position used to be that companies utilized intermediaries to create time and place utility, although there were conditions under which they could bypass middlemen and serve customers directly. The default position now appears to be that companies look first at going direct, although

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

many conditions exist under which it is beneficial to utilize specialized intermediaries.

Bibliography

Bailey, J.P. and Bakos, J.Y. (1997) An exploratory study of the emerging role of electronic intermediaries. *International Journal of Electronic Commerce*, **1** (3), 7–20.

Jallat, F. and Capek, M.J. (2001) Disintermediation in question: new economy, new networks, new middlemen. *Business Horizons*, **44** (2), 55–60.

Kauffman, R.J. and Chircu, A.M. (2000) Reintermediation strategies in business-to-business electronic commerce. *International Journal of Electronic Commerce*, **4** (4), 7–42.

point of difference and product differentiation

Rajendra Sisodia

In marketing, product differentiation refers to the process of taking what could otherwise be a commodity or homogeneous product and incorporating one or more useful and relevant points of difference into the overall product offering so that it offers more value to customers in an identified market segment. Product differentiation is one of the central concepts in marketing. The term was coined in 1933 by the American economist Edward Hastings Chamberlin, who also developed the theory of monopolistic competition. The objective of product differentiation is to create what advertisers term "a unique selling proposition."

For somebody with a marketing mind-set, there is no such thing as a commodity. Every product or service can be differentiated. This includes products intended for businesses as well as for consumers. Products need to be differentiated from those of competitors, as well as from a company's other products.

When products are differentiated, the offerings of different companies are no longer perfect substitutes, and customers cannot readily compare them with one another. This reduces the degree of competitive intensity in the market. In a sense, every company has a monopoly of its own product, since no one else makes exactly the same product. Differentiated products are not necessarily *better*, in an objective sense, than others; they are simply designed to appeal to a particular market niche.

Successful product differentiation moves products from competing primarily on price to competing on nonprice factors. The opportunities for product differentiation exist with the physical qualities of the product itself, as well as within the total package of benefits offered to the customer. These benefits could include elements such as differences in quality (usually accompanied by differences in price), differences in availability (such as time and location), packaging, service, installation, training, customization, brand image, accompanying products or services, the manner in which the product is delivered, and many other possibilities. Indeed, the opportunities for product differentiation are limited only by the imagination and creativity of marketers.

The importance of product differentiation actually increases when the core product offering is undifferentiated. Undifferentiated product offerings are essentially commodities, and if treated as such, do not require any marketing actions. One of the most important reasons the marketing function exists is to continuously decommoditize products, so that they offer differential value to different customers based on their unique needs and requirements.

It is helpful to think of a product as not a simple object but as a "complex cluster of value satisfactions," as Theodore Levitt put it. His "total product concept" consists of the generic product, the expected product (which includes the minimal attributes necessary in order to acquire any customers at all), the augmented product (which goes beyond the bare minimum requirements to include value-added elements that have different levels of appeal to different customers), and the potential product (which refers to additional ways to create value that remain possible, but have not yet been attempted).

Over time, many products that used to be sold as commodities have become highly differentiated offerings. Examples of foods include coffee, salt, flour, sugar, salt, oatmeal, bananas, chicken, pineapples, and many others. Relatively homogeneous services such as banking, insurance, credit cards, and air travel have become differentiated as well.

Product differentiation goes hand-in-hand with other central concepts in marketing such as MARKET SEGMENTATION AND TARGETING and marketing targeting and positioning (*see* POSITIONING ANALYSIS AND STRATEGIES). Product differentiation must be purposeful, in that it creates additional tangible and intangible value for the customers for whom the product is intended. Differentiating a product generally adds to the cost of producing it, and this cost must be more than offset by the additional value that is created for customers, and for which they are willing to pay a premium.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Bibliography

Dickson, P.R. and Ginter, J.L. (1987) Market segmentation, product differentiation and marketing strategy. *Journal of Marketing*, **51** (2), 1–10.

Levitt, T. (1980) Marketing success through differentiation – of anything. *Harvard Business Review*, **58** (1), 83–91.

Smith, W.R. (1956) Product differentiation and market segmentation as alternative marketing strategies. *Journal of Marketing*, **21** (1), 3–8.

positioning analysis and strategies

U. N. Umesh

Positioning is the act of designing an organization's offering and image so that it occupies a distinct and valued place in the target customer's mind relative to competitive offerings (Kerin and Peterson, 2007; Ries and Trout, 1982). The result of positioning correctly is the creation of a customer-focused value proposition (Kotler and Keller, 2006). Successful positioning requires that the similarities and dissimilarities between brands should be clearly defined and communicated in a target market for a particular category (*see* PRODUCT CATEGORY) or service. Further, the customer should find the position believable; otherwise the firm will never achieve its intended positioning in the minds of the customers who may even be turned off by the unrealistic positioning message. Given the specific nature of positioning, and the diversity of a market, it is increasingly difficult to create a brand position that appeals to everyone in all usage circumstances. Coca-Cola might have been one of the few firms that achieved universal appeal in the first half of the twentieth century using a very broad-based positioning for its single product brand. However, over the years, even Coca-Cola was forced to create separate formulations and unique positionings for those who preferred diet, orange or lemon-flavored product, and so on.

Positioning is of particular relevance in a global context (*see* GLOBAL MARKETING STRATEGY). Though the markets in industrialized countries are large, they are growing at a much slower pace as compared to those in developing countries. Not only are the markets in the two most populous countries in the world – India and China – growing at a rapid pace but numerous smaller countries with their widely different cultures are also growing at similar rates because of the advent of globalization and the unleashing of free markets. There are numerous global brands that appear to be successful in almost all markets (Helm, 2008). Quelch (2007) has suggested that having a single and consistent positioning worldwide with the company name and brand name being the same is the surest recipe for success – with a focus on a single product category. In contrast, Reibstein and Day (2004) suggest that it might be wise to have different levels of globalization providing some degree of local adaptation. As an example, food products should meet local cultural tastes in terms of branding.

An ideal positioning strategy is not necessarily achieved by aiming for the highest evaluative position in the minds of the customer. Rather, it is important that the positioning offered should match expectations of the customers in terms of the ideal attribute levels desired by the customers. Many firms are successful with a low-price/medium-quality brand image while other firms are equally successful with above-average price/high-quality brand image.

Mispositioning can occur when there is a mismatch between the ideal attribute levels desired by the customer and those offered by the firm. Usually, it is a positioning that the firm would like the customer to have, rather than what the customer actually desires. An example is the introduction of dry beer, which was rated as superb by brewers, where the customers had a hard time understanding how dryness of a beer could possibly quench the thirst of a parched throat.

Over time, the ideal attribute levels desired by the customers change. These changes could be due to advancement of technology, acceptance of new cultural values, changes in demographics, or improvements in competitive (*see* COMPETITIVE ANALYSIS) offerings. If there is no concomitant change in positioning, then the firm, product, or brand goes into the decline phase and can disappear altogether over time. Thus, there is a need to constantly evaluate and change positionings over time to match changes in customer requirements.

See also *global marketing strategy; product category*

Bibliography

Helm, B. (2008) *Best Global Brands*. Business Week (Sep 18).

Kerin, R.A. and Peterson, R.A. (2007) *Strategic Marketing Problems: Cases and Comments*, 11th edn, Pearson Prentice Hall, Upper Saddle.

Kotler, P. and Keller, K.L. (2006) *Marketing Management*, 12th edn, Pearson Prentice Hall, Upper Saddle.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Quelch, J.A. (2007) Why Global Brands Work, HBS Working Knowledge, October 17, 2007.

Reibstein, D. and Day, G. (2004) Managing brands in global markets, in *The INSEAD Wharton Alliance on Globalizing*, (eds H. Gatignon, R.E. Gunther and J.R. Kimberly), Cambridge University Press, 184–206.

Ries, A.l. and Trout, J. (1982) *Positioning: The Battle For Your Mind*, Warner books, New York.

pricing strategy

R. Venkatesh

Introduction

Marketing efforts revolve around the concept of value, especially as perceived by customers. Specific branches within marketing examine the magnitude of customers' perceived value of goods or services, the sources or drivers of value, and the ways and means of delivering better value to the customer and to the firm. Pricing and other marketing mix variables are strategic aids for a manager in the value creation and management process.

It is often noted that while product development, advertising, and distribution convey or deliver value to the customer, pricing emphasizes delivering value to the firm (i.e., extracting value from the customer). While this belief is largely correct, the interaction and interplay within the marketing mix cannot be overlooked. Thus, while one could consider changing prices in isolation of other factors, a better approach would be to implement price changes in tandem with other modifications to the value proposition (*see* A FRAMEWORK FOR CREATING VALUE PROPOSITIONS). As an example, a price increase on a Toyota Prius could be tied to commensurate improvements in the vehicle's fuel efficiency. Such an approach underscores what pricing truly is: an integral component of a firm's overall value proposition.

This article on pricing strategy is introductory in nature and has the following orientations. *First*, the article attempts to layout and connect the key elements in the pricing process. Pricing terms are introduced on a selective basis. *Second*, strategic pricing implications and recommendations are provided in almost every section. In contrast to strategies, tactical aspects of pricing (e.g., sales promotion tools such as "couponing") get limited attention. *Third*, some of the relatively more recent trends in or challenges to pricing are examined. Price taking as opposed to price setting (priceline.com) and auctions (eBay) are among the Internet-driven topics discussed. Throughout this article, goods and services are collectively referred to as *products*.

Among the strategic implications proposed in this article are the following:

- Value pricing is the desirable pricing approach. It overcomes the myopia of cost-plus and going-rate pricing without losing sight of costs and competing offerings.
- Product differentiation is a prerequisite for value pricing. Commodities drive prices down to marginal costs.
- A decreasing price is often more desirable than an increasing price path. High initial prices can signal superior product quality when other information is limited.
- While most price discrimination (PD) strategies are legal, the seller should not lose sight of price transparency and fairness.
- The Internet is opening up exciting and profitable avenues for pricing.

The emphasis on breadth and brevity means this article only serves to introduce the domain of pricing to the reader. Several books and numerous articles, from both the managerial and academic realms, are available for the interested reader. Some of these are listed in the concluding section. The ideas, concepts, and methods discussed in the following pages are drawn from or inspired by many sources. Per the policy of the editors, only a few references are cited explicitly. Kotler and Keller (2008), Nagle and Hogan (2005), and Rao (2009) are among the sources consulted most often during the writing of this article.

The remainder of this article has the following structure. Setting the initial price of a product is discussed in the section "Setting the Price." Managing this price over time and across segments are considered in the sections "Managing The Price Over The Life Cycle" and "Varying The Price Across Segments," respectively. The influence of competition on price setting is examined in the section "Anticipating And Responding To Competitors' Price Changes." The interrelationships among products in a firm's portfolio and their impact on the pricing of a particular product are discussed in the section "Pricing A Portfolio Of Complements Or Substitutes." The section "Pricing and the Internet" focuses on Internet-driven

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

pricing trends and challenges. A chapter summary appears in the "Conclusion" section.

Setting the Price

Price setting requires an appreciation of both the product and the market. For illustration, the success of repeat purchased goods, sometimes called *FMCG*s (*fast moving consumer goods*), depends on both trial and repurchase. Pricing practices that facilitate brand loyalty have merit. The seller of FMCGs can gain access to rich datasets such as scanner panel data as part of setting the prices. In contrast to FMCGs, durable goods are purchased with long inter-purchase times and typically after significant information search and consideration. On the market side, the nature of heterogeneity across customers in taste, brand preference, purchase incidence, and cognitive/monetary resources plays an important part in price setting. These types of factors will require adjustments to the price-setting process described below.

Eliciting or estimating individual value or aggregate demand is a prerequisite to price setting. The demand function, namely, the representation of the units of a product demanded at different price points, is often calibrated in a linear or curvilinear form as shown in Figure 1.

With two competing products 1 and 2, the linear demand function for 1 takes the form

$$Q_1 = a_0 + a_1 P_1 - a_2 P_2 \tag{1}$$

where the Ps and Q represent prices and quantity. Coefficient a_1 (or a_2) represents price sensitivity, namely, the change in quantity demanded of product 1 when P_1 (or P_2) is changed by one currency unit, holding all else constant. The constant elasticity demand function has a curvilinear form. Again, with two competing products 1 and 2, the demand function for 1 is

$$Q_1 = a_0 P_1^{\varepsilon_1} P_2^{\varepsilon_2} \tag{2}$$

where the Ps and Q continue to represent prices and quantity. The exponent ε_1 represents the self-price elasticity of demand of product 1 and ε_2 is cross-price elasticity of demand of product 1 with product 2. In either case, elasticity refers to the percentage change in quantity demanded for a 1% change in price. Typically, self-price elasticity is negative (i.e., an increase in a product's price causes its demand to decrease) whereas cross-price elasticity is positive (i.e., an increase in a competitor's price causes the focal product's demand to increase). The choice between the above two functional forms can be on the basis of goodness of fit between historical price and demand data.

When the marginal cost of product 1 is c_1, and given the price of the competing product, the profit-maximizing seller can set the price P_1 such that

$$\underset{P_1}{\text{Max}} \prod_1 = (P_1 - c_1) \times Q_1 \tag{3}$$

This again assumes that the seller seeks profit maximization in each time period. (The competitor is also strategic and that requires understanding the best response of each firm and the resulting equilibrium. This is taken up in the section "Anticipating And Responding To Competitors' Price Changes.") Other objectives such as setting the best price to sell a targeted quantity or maximizing the net present value of a stream of profit flows can all be achieved with a proper understanding of demand.

The above approach to price setting assumes a clear knowledge of (i) customers' valuations or purchasing behavior, (ii) firm's cost structure, (iii) competitors' prices and intentions, and (iv) product interdependencies. Such a comprehensive knowledge may not always exist. Alternative approaches are discussed below.

Cost-plus (or markup) pricing and going-rate (or competitor-based) pricing are more prevalent than what sound strategy might dictate. In

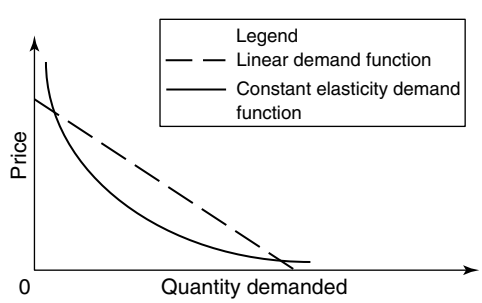


Figure 1 Two popular forms of the demand function.

Table 1 A simple illustration of value pricing.

(This is a fictitious example)		
A chemical plant uses 400 O-rings each year to seal valves. Rings are priced at \$8.00 each and must be changed every two months in a shutdown process that costs \$8000 (including the labor cost of replacing the rings). Our firm has developed a new O-ring that has twice the resistance to corrosion. The marginal cost to our firm of the new O-ring is \$5. There are no other alternative O-rings in the marketplace. What is the economic value advantage of each new O-ring?		
Let the price of each new O-ring be *P*		
Description	**Regular O-Ring**	**New O-Ring**
Number of rings needed	400	200
Purchase price to the customer	\$3 200	200 × P
Shutdown costs	\$48 000	\$24 000
Total costs to the customer	\$51 200	\$24 000 + 200P

Economic value advantage of each new O-ring $= \frac{\$51200-\$24000}{200} = \$136$. Our firm would not want to charge a price lower than \$5 for the new O-ring as that would yield a negative unit contribution margin. \$136 is the unit price at which the customer would be indifferent between the regular and the new O-rings. The actual price is likely to be between these extremes and based on other considerations such as the likelihood of new entrants at different candidate prices.

cost-plus pricing, the firm adds a markup to the unit cost of the product. Going-rate pricing considers the price charged by the leading (or closest) competitor, with some upward or downward adjustment based on whether the firm is a price leader or follower. Both of these approaches are easy to implement yet deeply flawed. The ease arises from the firm's knowledge of its own cost structure and/or the prices of competing offerings. The problems arise from taking a very passive and myopic view of pricing and marketing, namely, that the firm takes its cost and competitors' prices as a given and is not actively seeking to differentiate its offering and enrich its value proposition.

Value pricing is arguably a superior alternative as it balances firm's costs, value of competing offerings, and the incremental value of the product under consideration. Thus

$$\text{Perceived value to the customer} = \text{Reference value} + \text{Differentiation value} \tag{4}$$

The reference value is the net value of the closest substitute. The differentiation value represents whatever value the focal product offers above and beyond the reference product. This assessment of value considers all the sources of benefits and costs (tangible and intangible) to the customer. Table 1 contains a simple illustration with higher value equated to lower costs.

In the table, the term *economic value advantage* (*EVA*) represents the differentiation value without factoring in the price of the new offering – the decision variable in this exercise. The portion of the EVA to extract in the form of price will depend on the firm's strategic priorities.

An added advantage of value pricing is that it can be used even to price radical innovations (e.g., the MRI machine, when it was first introduced). Such innovations have weak substitutes (e.g., CT scan in lieu of MRI). Thus, while the technology is new, the benefit can be quantified on the basis of what the available substitute does or does not offer. Techniques such as conjoint analysis augment value pricing by looking at attribute-level value drivers.

The term *reservation price*, defined as the maximum price a customer is willing to pay for one unit of a product, is a currency metric of the customer's perceived value. The difference between a customer's reservation price and the actual price of the product represents his or her

surplus. When a product enjoys a monopoly, the prospective customer is expected to buy the product if she/he expects a nonnegative surplus. When competing products are available, the customer is expected to choose the product that offers the highest nonnegative surplus.

Managing the Price Over the Life Cycle

The STAGES OF THE PRODUCT LIFE CYCLE is a potent marketing paradigm to conceptualize the evolution of alternative products and draw strategic implications over time. A successful product goes through four stages in its life cycle – introduction, growth, maturity, and decline. Sales during maturity come primarily from product replacement. The length of each stage is partly a function of the characteristics of the product and of potential consumers such as their attitudes toward risk and the level of word of mouth among themselves. Seen from a pricing standpoint, a firm can "manage" the adoption by choosing from among alternative price trajectories. Such variation of price over time is called *dynamic pricing*.

Skimming and penetration pricing are two alternatives to a more neutral form of pricing. With skimming, the firm starts by charging a high price and then gradually decreases it over time. This could be very profitable for durables: the decreasing price path could extract the most value from customers, assuming that customers will not delay their purchases inordinately in anticipation of lower prices in the future. With penetration pricing, the product is initially priced aggressively low to encourage speedy adoption. How the price is changed later is a function of other variables to be discussed. The choice in favor of skimming or penetration pricing in relation to neutral pricing depends on variety of firm-, competitor-, and customer-side factors.

Skimming is a more desirable under the following conditions:

- *Firm-side factors*: When a firm has a wide product line (e.g., Intel, with its range of microprocessors), a new product can be priced high enough so as not to cannibalize other offerings in the line. Also, a firm's capacity constraint naturally lends itself to skimming.
- *Competitor-side factors*: Skimming is not only a more time-intensive strategy (as the market is built gradually, over time) but also the higher price might signal low barriers for entry. Therefore, skimming is more desirable only when the firm has a clear technological advantage (e.g., Apple with iPhone) or patent protection to milk the market over time.
- *Customer-side factors*: Greater heterogeneity in customers' valuations and the willingness within segments to pay significant premiums for their "ideal" products favors skimming. It is desirable that price skimming is not excessive and is not seen as unfair. This occurred with HIV Aids medication, for example, when the high initial price of the drug cocktail (over $10 000 per patient per year) caused a furor from the developing world. Global prices quickly dropped to under a dollar a day per patient.

Conditions opposite to the ones noted above for skimming are likely to favor penetration pricing. The following additional factors also point to a low penetration price:

- *Firm-side factors*: Penetration pricing is desirable when a firm expects significant gains from economies of scale. Also, for experience goods such as software, an extreme penetration pricing approach of offering a free initial version to gain a foothold and a later upgrade at a higher price may be optimal.
- *Competitor-side factor*: When two competing systems are in a standards war, the one that establishes a larger user base would arguably emerge a winner. Penetration pricing could facilitate this. The lower price of Toshiba's HD DVD system during its losing battle with Sony's BluRay DVD format stands out as one of the exceptions.
- *Customer-side factor*: For products with network externalities, the value to individual users is a function of the number of other users of the product. A low initial price to build the network followed by a higher price

later in the life cycle has been suggested as an optimal strategy.

Every-day-low-price (EDLP) and high–low (Hi–Lo) pricing are also forms of dynamic pricing. As these are not tied to the product life cycle paradigm and are of a tactical nature, these are not discussed further in this article.

Varying the Price Across Segments

While a firm should strategically arrive at a price for its core segment of users, a one-price-fits-all approach toward the entire market is seldom the most profitable. Segments differ in their preference intensities for the product, in their income levels and price sensitivity, quantity of consumption, and the like. Price discrimination (PD) is a desirable strategy in such instances and is said to occur "when a company sells a [good] or service at two or more prices that do not reflect a proportional difference in costs" (Kotler and Keller, 2008).

While three degrees – or types – of PD have been identified, these differ in the ease of implementation. Under first-degree PD, the seller charges each customer a price equal to her reservation price. Of course, only those customers with reservation prices higher than the product's marginal cost are targeted. When possible, this type of PD should be the most profitable for the seller as the entire surplus is extracted. Two common hurdles to first-degree PD are the seller's lack of awareness of each customer's reservation price and the inability to charge different prices without upsetting customers' perceptions of fairness. Auctions are one way of implementing first-degree PD. If bids are invited and one or more units of the product are sold to the highest bidder(s), the seller is arguably extracting the full surplus from at least the winning bidder(s). Singapore Airlines' auction of seats on the first commercial flight of the Airbus A380 Super Jumbo, between Singapore and Sydney, can be called first-degree PD. Here, interested fliers submitted bids on eBay and tickets priced on par with the respective bids were awarded to the highest bidders till all seats were sold out.

Second-degree PD occurs when the seller offers different prices to different buyers based on their purchase quantity. An objective of this form of PD is to encourage buyers to purchase in larger volume. The law of diminishing marginal returns is an associated rationale for second-degree PD. As a second or third unit gives less utility to the consumer than the first, discounts are warranted on these latter units to facilitate purchase. In an excellent review article, Dolan (1987) contrasts the uniform price schedule (i.e., no PD, as each unit is sold at the same price) from n-part and n-block price tariffs. In a simple two-part tariff ($n = 2$), each buyer pays a fixed fee plus a variable rate tied to the number of units purchased. This tariff structure reduces the unit price for a large purchase because the fixed fee is apportioned over a larger number of units. Telephone calling plans and some types of software licenses come under this category. In a two-block tariff, a price P_1 is charged for each for the first N units; a price $P_2(< P1)$ is charged for each unit beyond N. The three-part or three-block tariff should price discriminate at least as well as a two-part or two-block tariff. The issue to guard against is customer confusion from too fine a structure.

Third-degree PD occurs when the distinctive aspects of customers (e.g., senior citizen vs. young adult) or shopping occasions (e.g., matinee vs. regular show) form the basis for charging different prices. Unlike second-degree PD, the quantity of purchase is not the discriminating variable. Unlike first-degree PD, the differences in reservation prices are not known for sure. Rather the expectation is that some groups of buyers may have lower (or higher) reservation prices. Senior citizens and students who probably do not have full-time jobs are likely to have more available time or less disposable income, thereby making themselves appropriate targets for price discounts. For services (e.g., flights), that are cannot be inventoried unlike physical goods, and for offerings subject to capacity constraints (e.g., concerts), third-degree PD may help in managing demand (by shifting part of the peak load to off-peak times).

Although there are legal restrictions on some forms of PD (e.g., some types of quantity discounts must be backed by evidence of associated cost savings to the seller), most forms of PD are legal and hence widely used. Constraints to PD may arise from the buyer side. PD will be

rendered ineffective or inappropriate if, say, it is perceived as blatantly unfair or if those buying the products at a lower price sell these to others who might not have qualified for the discount (a practice called *arbitrage*). In most situations, PD actually increases customer welfare by making products available at lower prices to specific groups and increasing consumption.

Anticipating and Responding to Competitors' Price Changes

The role of competition can never be overemphasized. Be it an entrenched player or a capable new entrant, competitors are "hurdles" in a firm's pursuit of higher profits. By the same token, healthy competition results in better value for customers by fostering innovation, keeping prices under check, and tailoring marketing practices to suit the needs of specific segments.

Work on competitive strategy (Porter, 2005) and game theory (Dixit and Nalebuff, 1993) provides ways to understand competitors' incentives and motives, and take preventive, protective, or reactive measures (*see* COMPETITIVE ANALYSIS; COMPETITOR ANALYSIS). Table 2 contains a simple illustration of game theoretic thinking that has the following elements: (i) the players are aware of each other's strategic options and the corresponding payoffs; (ii) each player recognizes that the competitor is not static but behaves strategically; (iii) each player acts in one's own best interests knowing that the other player does the same; that is, benefiting oneself rather than hurting the other is the overriding motive.

The illustration underscores that while higher prices from the two players would have provided higher payoffs for both, thinking strategically about each other in a noncollusive manner leads to a lowering of prices. The example shows how the free market can work to the customer's advantage.

There is also a limitation in the above approach. The illustration assumes that price is the only variable at the firms' disposal at this stage of the game. A broader marketing view of this problem could consider nonprice ways of mitigating the intensity of competition. Echoing this thought, Porter (1998) notes that there are essentially two competitive strategies: gaining cost advantage or pursuing differentiation. The

Table 2 Pricing under competition: a simple illustration.

(The numbers are made up)

Ford and General Motors (GM) are the two dominant competitors in the B2B market to sell cars to rental car companies. This business segment is marked by aggressive price discounting unlike the end user market. The discounts are usually either 15% or 20% off the maximum suggested retail price (MSRP) in the retail market. The players do not collude in any way; they play the Leader (GM)–Follower (Ford) game. That is, GM sets its discount rate first. Ford sees it and immediately comes up with its best response. The payoff matrix is given below.

		GM's Discount	
		15% off MSRP	20% off MSRP
Ford's discount	15% off MSRP	\$360M, \$640M[a]	\$180M, \$410M
	20% off MSRP	\$420M, \$240M	\$210M, \$350M

[a] Means that Ford's payoff is \$360 million and GM's payoff is \$640 million. What discount levels are Ford and GM likely to offer at the equilibrium? Why? In this example, the equilibrium corresponds to GM offering a 20% discount and Ford responding to it with a 20% discount. The reasoning is as follows: no matter what discount GM offers (15% or 20%), Ford's best response is a 20% discount. This is because (i) when GM offers a 15% discount, Ford makes \$420M with a 20% discount instead of \$360M with a 15% discount, and (ii) when GM offers a 20% discount, Ford makes \$30M more for itself with a 20% discount relative to a 15% discount. Knowing that Ford's best response is a 20% discount, GM will offer a 20% discount as it will make \$350M (i.e., \$110M more than with a 15% discount). With both firms offering a 20% discount, neither will have a unilateral incentive to deviate, and this discount combination will represent a Nash equilibrium.

firm with the lowest cost structure in its industry (e.g., Southwest Airlines) has a greater ability to fight a price war. Everyone else is better off pursuing differentiation.

Price wars represent the uglier side of pricing, the competing sides emerging losers in the fight. Rao, Bergen, and Davis (2000, RBD) point out that from the firm's side, the best approach to a price war is often not to get into one. In the guidelines below on how to avoid a price war or mitigate the effects of one, those from article by the Rao *et al*. are tagged RBD:

- *If a competitor attacks you in your backyard, signal your intent by attacking the competitor in its backyard*: The example of Michelin, with its European stronghold, initiating a price war against Goodyear in its prime US market is well known. Goodyear responded by cutting prices in Europe, signaling to Michelin that it stood to lose considerably. The price war subsided.
- *Reveal your cost advantage to your competitor (RBD)*: The authors make the point that revealing this advantage will make the competitor aware that you have more room to cut prices and still make a positive contribution.
- *Buy excess capacity in your industry and shutter the plants*: Industries with excess, idle capacity are prime candidates for price wars. Try to reduce this excess.
- *Introduce a fighting brand (RBD)*: Offerings such as MCI Worldcom's 10-10-220, Kodak's Funtime film, and Goodyear-owned Kelly Springfield tires are examples of high-quality firms introducing lower-end fighting brands to handle price competition without exposing the flagship offerings to the fight.

The above points underscore that pricing strategies have more power when firms are willing to tap into their broader marketing options.

Pricing a Portfolio of Complements or Substitutes

There are few single-product firms even though many a firm may have a single blockbuster. Products in a firm's portfolio could be (partial) substitutes of each other when they are part of a product line (e.g., the line of soft drinks offered by Coca Cola). When a firm is in multiple categories, the products could be complements of each other (e.g., Samsung's DVD player is a complement of the firm's TV lines). This section makes the point that pricing strategies are more potent and meaningful when the interrelationships among the offerings are suitably leveraged. The degree of substitutability or complementarity among the offerings has strategic significance.

Cross-price elasticity is one measure of the degree of interrelationship. Two products are complements of each other when their mutual cross-price elasticities are negative; that is, an increase in the price of one causes a decrease in demand for the other. Cross-price elasticities between two substitutes are positive. Another measure is by how much the reservation price of a bundle is greater than (or less than) the sum of the stand-alone reservation prices of the individual products. For a bundle of complements (or substitutes), the reservation price of the bundle is greater than (or less than) the sum of the individual reservation prices.

Consider a product line of moderate to weak substitutes. (Having two strong substitutes in a firm's product line is arguably suboptimal.) There is evidence to suggest that with a wider product line, the price of each product is typically higher relative to the case when that product alone is offered. The rationale is that with several offerings, each can be targeted at the segment that it is best suited for. The match with the segment's ideal point justifies charging a higher price. A product line is also justified as a way to keep up with an evolving market. Considering again the example of the line of Intel microprocessors for personal computers, the firm cuts prices of existing microprocessors as newer ones are introduced at high price levels. Discounting the more established processors is a bulwark against competitors. Thus, the line remains very profitable while also protecting the firm's market share.

For a portfolio of complements, BUNDLING is pertinent. *Bundling* is the strategy of offering of two or more products as a specially priced package. Bundling encourages customers to

try the complementary products jointly. The synergy among complements means customers see higher value and the firm can tap into some of this value enhancement. Three bundling strategies must be distinguished. Under pure components, the products are offered separately but not in bundled form. Pure bundling occurs when the products are offered only as a bundle but not in stand-alone form. Mixed bundling is the more general strategy of offering both the individual products and the bundle. This strategy helps achieve second-degree PD by offering the discounted bundle at the segment that cares for all of the component products, and charging price premiums on the individual offerings.

The choice among alternative bundling strategies is further clarified below:

- *Firm-side factors*: Pure bundling is favored when it can significantly reduce inventory holding costs (e.g., a car model offered in a few popular trims instead of many attribute combinations) and/or sorting costs (e.g., raw diamonds, sold as bundles by DeBeers).
- *Competitor-side factors*: Although pure bundling could deter competition in some instances, the pure components strategy is desirable when customers' strongly preferred bundle is a mix of offerings from rival firms (e.g., Dell PC with HP laser printer).
- *Customer-side factors*: Mixed bundling, the equivalent of a wider product line, is desirable when customers' preferences are very heterogeneous. Pure bundling is desirable when it is harder for a customer to identify the cause when a bundle performs poorly. Otis encourages customers to opt for its own service plan because, if the product breaks down because of poor maintenance by another service provider, customers may still blame Otis for the problem.

Pricing and the Internet

The Internet provides new and exciting pricing opportunities and challenges (*see* E-COMMERCE AND INTERNET MARKETING). The challenges come in several forms: (i) prospective consumers can easily and quickly compare prices and product specifications; (ii) blogs and social networking websites disseminate word of mouth with speed and on a large scale; (iii) products that can be digitized, such as movies and music, can also be pirated; (iv) magazines and newspapers are threatened by the large amount of free news content. Yet the Internet has spawned entirely new businesses and business models that depend on clever pricing strategies. A core component of any business plan is the revenue model – a clarification of how the firm expects to price and make money.

Two online business models with innovative pricing arrangements are discussed here:

- *Name your own price*: The emergence of Priceline.com (among others) has meant that the Internet has made pricing buyer driven rather than seller driven. The buyer submits a bid for a product such as an airline ticket between two cities, and the seller decides whether to accept or reject the offer. Thus the seller becomes a price taker instead of being a price setter. Frivolous bidding is limited by imposing restrictions on rebidding. The Priceline model cleverly combines pricing and product design. While a bidder quotes a price, it is up to the seller to select the airline and specific schedule.
- *Auctions*: eBay has made auctions a household name. By establishing a trading platform, trading rules, and mechanisms, eBay has dramatically increased the number of sellers and the way they sell. Network externalities are integral to auctions: an auction site is more attractive to a buyer (or seller) if it attracts more sellers (or buyers) than a competing site. By facilitating these connections, eBay is increasing revenues to sellers while simultaneously delivering greater value to the buyers.

The Internet's role in pricing and marketing is just beginning.

Conclusion

This article has sought to provide a broad yet cursory overview of the strategy of pricing. It has introduced and discussed pricing terms that are frequently invoked. The article has also

emphasized that an effective pricing strategy often draws on other elements of the marketing mix. For setting and managing the price of a product many elements such as the segments in the marketplace, the intentions and actions of competitors, the interrelationships with other products in the portfolio, and the power of new technology and channels such as the Internet have to be considered. A number of strategic implications have been discussed. Yet this article is incomplete. The following paragraphs suggest the notable omissions in this article and suggest some ways of gaining a deeper understanding of pricing.

The pricing implications of consumer psychology and behavior (*see* CONSUMER BEHAVIOR ANALYSIS) have barely been noted in this article. Prospect theory, for example, focuses on how customers respond differently to losses and gains. The price of a product represents a cost to the customer. Rebates and discounts can be seen as gains. Related implications of prospect theory can guide how prices and discounts should be framed to have more favorable impact. This article has barely scratched the surface on the topic of channel pricing. As intermediaries, channel members can cause an increase in product prices even while providing important services. The success of emerging channel structures (*see* MARKETING CHANNEL STRATEGY) such as partial forward integration and bricks-and-clicks outlets depends in part on appropriate pricing. The intertwined concepts of customer lifetime value and customer relationship management require taking a relationship perspective to marketing mix efforts, including pricing. Pricing methodologies span the spectrum from measurement of customers' valuations to multiproduct pricing under competition. These and other topics are missing in this article.

A short list of references at the end of this article can offer the reader a more in-depth view of pricing. Most of the references are aimed at managers. Two references for the theoretically driven reader are Rao (2009) and Wilson (1993). Marketing journals such as the *Journal of Consumer Research*, *Journal of Marketing*, *Journal of Marketing Research*, and *Marketing Science* are recommended for rigorous studies on pricing. The *Harvard Business Review* is one of the several publications taking a managerial view of emerging topics. Nuances in pricing and related marketing of digital products are discussed in Shapiro and Varian (1998). Pricing strategy is an exciting and important aspect of global marketing (*see* INTERNATIONAL PRICING OBJECTIVES AND STRATEGIES). Mahajan and Banga (2005) provide an insightful reference on the opportunities and challenges of marketing to the 86% of the world's population that lives in the developing world. The above references can facilitate a greater appreciation of the fascinating topic of pricing.

Acknowledgments

The author thanks Rabikar Chatterjee and Editors Roger Kerin and Robert Peterson for their helpful comments on an earlier version of this manuscript.

Bibliography

Dixit, A.K. and Nalebuff, B.J. (1993) *Thinking Strategically: The Competitive Edge in Business, Politics, and Everyday Life*, W.W. Norton & Company, New York.

Dolan, R.J. (1987) Quantity discounts: managerial issues and research opportunities. *Marketing Science*, **6** (1), 1–22.

Dolan, R.J. and Simon, H. (1996) *Power Pricing: How Managing Price Transforms the Bottom Line*, The Free Press, New York.

Kotler, P. and Keller, K.L. (2008) *Marketing Management*, Pearson Education, Inc., Upper Saddle River.

Mahajan, V. and Banga, K. (2005) *The 86% Solution: How to Succeed in the Biggest Market Opportunity of the 21st Century*, Wharton School Publishing, Philadelphia.

Nagle, T.T. and Hogan, J. (2005) *The Strategy and Tactics of Pricing: A Guide to Growing More Profitably*, Pearson Education, Inc., Upper Saddle River.

Porter, M.E. (1998) *Competitive Advantage: Creating and Sustaining Superior Performance*, The Free Press, New York.

Rao, V.R. (2009) *Handbook of Pricing Research in Marketing*, Edward Elgar, Northampton.

Rao, A.R., Bergen, M.E., and Davis, S. (2000) How to fight a price war? *Harvard Business Review*, (March-April), 107–116.

Shapiro, C. and Varian, H.R. (1998) *Information Rules: A Strategic Guide to the Network Economy*, Harvard Business School Press, Boston.

Wilson, R.B. (1993) *Nonlinear Pricing*, Oxford University Press, New York.

product category

Jaeseok Jeong

The term *product category* is given to a group of similar products that fulfill a particular type of need and can be interchangeably used with product class. A group of products or brands in the same product category are homogeneous or generally considered as substitutes for each other. For example, Levi's and Gap belong to the jeans category and one can be a good substitute for the other. They both can meet consumers' particular need for jeans.

A products is generally identified in terms of its product category, which can be defined as the classification to which the product is assigned. For example, Special K is the breakfast cereal category and MacBook Air is in the notebook computer category. There are variations among products within a category in terms of physical characteristics (e.g., color, shape, attribute, or fragrance) or brand names. A single manufacturer may produce several brands within the same product category in order to appeal to different segments of the market. For example, Apple has multiple notebook computer brands such as MacBook, MacBook Pro, and MacBook Air in the same notebook category and markets these brands to different segments of its target customers to maximize its profits.

The product category can be narrow or broad depending on how substitutable the various products are. For instance, a narrow product category for a transportation vehicle would include motorcycles, cars, pickups, and vans, whereas a broad product category would include airplanes, trains, and ships. Another example is that a narrow product category for breakfast meats might be sausage, bacon, and ham and a broad product category would include all other meats and meat substitutes occasionally sold for breakfast use.

The identification of product category can sometimes be ambiguous and even arbitrary and this can be an issue for marketing managers in measuring the market performance and understanding the market structure and competition for their own product(s). If you are a marketing manager in a company producing breakfast meats, would you think that you are in the breakfast meats category and your competitors are all other (or similar) brands in the same product category or would you consider all other meats or meat substitutes brands including breakfast meats as competitors? In other words, are you in the breakfast meats category or in the meats and meat substitutes category? Your marketing strategy would be dramatically different depending on how you see your product category.

This inherent ambiguity and arbitrariness on the product category have led to debates on how it should be defined to better comprehend the market and its structure (Hanan, 1974; Moran, 1973). Day, Shocker, and Srivastava (1979) proposed that marketers focus on the benefits that products provide instead of the attributes or features that products have in identifying the product category. They recommended that the product category be defined as "the set of products judged to be substitutes within those usage situations in which similar patterns of benefits are sought, and the customers for whom such usages are relevant." By doing so, marketers could have a better understanding of the market and its structure and arrive at effective criteria of measuring the market performance and competition.

A product category tends to have a STAGES OF THE PRODUCT LIFE CYCLE that is longer than that of any one product within the product category alone unless a particular brand that was the FIRST-MOVER (PIONEER) ADVANTAGE in a market was also the last one.

Category development index (CDI), a market's category sales percentage divided by the total population percentage of that market multiplied by 100, and brand development index (BDI), a market's brand sales percentage divided by the total population percentage of that market multiplied by 100, are useful to determine market potential for a brand or product. Any number over 100 for the CDI and BDI is considered good but comparing the BDI to the CDI offers the most insight.

Primary advertising, also known as *generic advertising*, is the advertising intended to create demand for a product category of a product rather than a specific brand and to enhance the image of the entire product category involved. The campaign, "Got Milk" by California Milk

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Processor Board would be a good example of primary advertising. Primary demand is the demand for a general product category, as contrasted with selective demand, the demand of a specific brand marketed by a firm.

Bibliography

Day, G.S., Shocker, A.D. and Srivastava, R.K. (1979) Customer-oriented approaches to identifying product-markets. *Journal of Marketing*, **43** (4), 8–19.

Golder, P.N. and Tellis, G.J. (1993) Pioneer advantage: marketing logic or marketing legend. *Journal of Marketing Research*, **30** (2), 158–170.

Hanan, M. (1974) Reorganize your company around its markets. *Harvard Business Review*, **79**, 63–74.

Kotler, P. and Keller, K.L. (2009) *Marketing Management*, 13th edn, Prentice Hall, Upper Saddle River.

Moriarty, S., Mitchell, N. and Wells, W. (2009) *Advertising: Principles and Practice*, 8th edn, Prentice Hall, Upper Saddle River.

Moran, W.R. (1973) Why new products fail. *Journal of Advertising Research*, **13**, 5–13.

Plummer, J.T. (1974) The concept and application of life style Segmentation. *Journal of Marketing*, **38** (1), 33–37.

Raju, J.S. (1992) The effect of price promotions on variability in product category sales. *Marketing Science*, **11** (3), 207–220.

Ratneshwar, S. and Shocker, A.D. (1991) Substitution in use and the role of usage context in product category structures. *Journal of Marketing Research*, **28** (3), 281–295.

stages of the product life cycle

Jaeseok Jeong

The product life cycle (PLC) is a biological metaphor that traces the stages of a product's acceptance from its introduction (birth) to its decline (death). It can be divided into several stages characterized by the revenue generated by the product. The product revenue and profits can be plotted as a function of the life cycle stages as shown in Figure 1.

The concept of PLC may apply to a brand or to an entire PRODUCT CATEGORY. Its duration may be as short as a few months for a fad item or century or more for product categories such as the cola beverage and the gasoline-powered automobile. As the product progresses through its life cycle, changes in the MARKETING MIX are usually required in order to adjust to the evolving challenges and opportunities.

Introduction Stage

When the product is introduced, sales will be low until customers become aware of the product and its benefits. During the introduction stage, the firm aims to build awareness and develop a market for the product. The impact on the marketing mix is as follows:

Product – There is one or few products, relatively undifferentiated in the market. Product branding and quality level is established.

Price – Pricing (*see* PRICING STRATEGY) may be high, assuming a skim pricing strategy for a high profit margin as the early adopters buy the product and the firm seeks to recoup development costs quickly. In some cases, a penetration pricing strategy is used and introductory prices are set low to gain market share rapidly.

Distribution – Distribution is selective and scattered as the firm commences implementation of the distribution plan.

Promotion – Promotion is aimed at building brand awareness. Samples or trial incentives may be directed toward early adopters. The introductory promotion is also intended to convince potential resellers to carry the product.

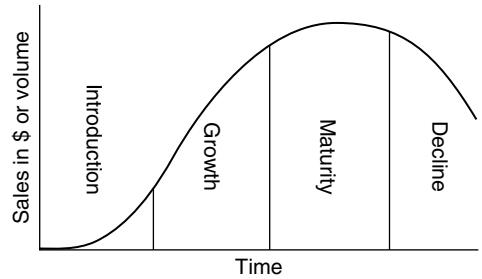


Figure 1 Several stages of product life cycle (PLC) can be characterized by the revenue over time.

Growth Stage

During the growth stage, sales increase as more customers become aware of the product and its benefits and additional market segments are targeted. The firm seeks to build brand preference and increase the market share.

Product – Product quality is maintained and additional features and support services may be added.

Price – Pricing (*see* PRICING STRATEGY) is maintained at a high level if demand is high, or reduced to capture additional customers.

Distribution – Distribution becomes more intensive. Trade discounts are minimal if resellers show a strong interest in the product.

Promotion – Advertising is increased to build brand preference.

Maturity Stage

The maturity stage is the most profitable. While sales continue to increase into this stage, they do so at a slower pace. Competition may appear with similar products. The primary objective at this point is to defend market share while maximizing profits and extend the PLC.

Product – Modifications are made and features are added in order to differentiate the product from competing products that may have been introduced.

Price – Price reductions are expected in response to competition while avoiding a price war.

Distribution – Distribution becomes more intensive and incentives to resellers may be offered to avoid losing shelf space.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Promotion – Promotion emphasizes product differentiation and building of brand loyalty.

Decline Stage

Eventually, sales begin to decline as the market becomes saturated. At this stage of the PLC, the firm has several options. First, the firm maintains the product and rejuvenates it by adding new features and finding new uses. Second, the firm harvests the product by reducing costs and continuing to offer it, possibly, to a loyal niche segment. Lastly, the firm discontinues the product when no more profit can be made or there is a successor product. The MARKETING MIX decisions in the decline stage depend on the selected strategy. For example, the product may be changed if it is being rejuvenated, or left unchanged if it is being harvested. The price may be maintained if the product is harvested, or reduced drastically if liquidated.

The term *life cycle* implies a well-defined cycle as observed in living organisms, but products do not have such a predictable life, and the specific life cycle curves followed by different products vary substantially (Dhalla and Sonia, 1976; Day, 1981). Hence the life-cycle concept is not well suited for the forecasting of individual product sales. Furthermore, critics have argued that the PLC may be self-fulfilling. For example, if sales peak and then decline, marketing managers may conclude that the product is in the decline stages and therefore cut the promotion budget, thus precipitating a further decline.

Nonetheless, the PLC concept helps marketing managers to plan alternate marketing strategies to address the challenges that their products are likely to face. It also is useful for monitoring sales results over time and comparing them to those products having a similar life cycle.

Bibliography

Anderson, C.R. and Zeithaml, C.P. (1984) Stage of the product life cycle, business strategy, and business performance. *Academy of Management Journal*, **27** (1), 5–24.

Day, G.S. (1981) The product life cycle: analysis and applications issues. *Journal of Marketing*, **45** (4), 60–67.

Dhalla, N.K. and Sonia, Y. (1976) Forget the product life cycle concept!. *Harvard Business Review*, **54** (1), 101–112.

Hofer, C.W. (1975) Toward a contingency theory of business strategy. *Academy of Management Journal*, **18** (4), 784–810.

Klepper, S. (1996) Entry, exit, growth, and Innovation over the product life cycle. *American Economic Review*, **86** (3), 562–583.

Kotler, P. and Keller, K.L. (2009). *Marketing Management*, 13th edn, Prentice Hall, Upper Saddle River.

Lambkin, M. and Day, G.S. (1989) Evolutionary processes in competitive markets: beyond the product life cycle. *Journal of Marketing*, **53** (3), 4–20.

Levitt, T. (1965) Exploit the product life cycle. *Harvard Business Review*, **43**, 81–94.

Thietart, R.A. and Vivas, R. (1984) An empirical investigation of success strategies for business along the product life cycle. *Management Science*, **30** (12), 1405–1423.

push and pull marketing strategies

Deanne Brocato

A push/pull strategy refers to the way in which information and products move between consumers and a manufacturer. Specifically in marketing, this strategy refers to the focus of the promotional efforts used to sell a good or service. In a push strategy, suppliers "push" their goods toward consumers, and in a pull strategy consumers "pull" information or goods that are suitable for their needs. A *pull strategy* is defined as a promotion strategy (*see* INTEGRATED MARKETING COMMUNICATION STRATEGY) focused on consumers rather than on members of the marketing channel in order to facilitate the flow of a good or service from a manufacturer to end-users (consumers). Conversely, a *push strategy* is defined as a promotion strategy focused on marketing intermediaries (wholesalers and retailers) rather than on consumers in order to facilitate the flow of a good or service from a manufacturer to consumers.

Push Strategy

A push strategy is a promotional strategy that "pushes" a product to the consumer through a distribution channel (intermediaries). This strategy focuses marketing activities on the use of aggressive selling (*see* SALES FORCE STRATEGY) and trade advertising (*see* INTEGRATED MARKETING COMMUNICATION STRATEGY) in order to convince intermediaries to carry, promote, and sell the product to consumers. In order to sell products to consumers, retailers often use advertising, point of purchase displays, and other promotions to convince consumers to buy the products. The advantage of this strategy is that it requires less manufacturer capital for advertising and promotional campaigns since awareness of a product is created by having it available for purchase in retail stores. Disadvantages of this strategy include inventory that retailers need to carry. In addition, a push strategy usually involves incentives to the distribution channel that increase marketing costs for the manufacturer. A push strategy is appropriate for products with low brand loyalty, where brand choice is made in the retail store, and impulse purchases (Kotler, 2003). As an example of a push strategy, the Jamaican Tourist Board markets and promotes its island to and through travel and tour agencies. In return, these agencies market directly to consumers (Lamb, Hair, and McDaniel, 2006). This is considered a push strategy since promotional efforts are aimed at channel intermediaries (travel and tour agencies) instead of directly at consumers.

Pull Strategy

A pull strategy is a promotional strategy that focuses on building up consumers' demand for a particular product. ADVERTISING MEDIA SELECTION AND PLANNING and direct-to-consumer promotions are used as a way to motivate consumers to search for and seek out a certain brand of product. Instead of retailers stocking products because of the incentives from a manufacturer, retailers purchase products from the manufacturer because of increasing demand from consumers. The manufacturer uses advertising, sampling, couponing, and other consumer promotions to encourage consumers to ask intermediaries for the product, thus encouraging the intermediaries to order the product. A pull strategy requires highly visible brands that can be developed through advertising campaigns, publicity, or word-of-mouth referral. The advantage of this strategy is that there is less retailer inventory since consumers' demand drives the retailer to purchase from the manufacturer. The disadvantage of this strategy involves the high cost of advertising campaigns and publicity efforts to create a highly visible brand that will stimulate consumers' demand in the marketplace. A pull strategy is appropriate for brands that have high levels of loyalty, when people understand the benefit differences between brands, and consumers can choose a brand prior to entering a retail store (Kotler, 2003). As an example of a pull strategy, a popular liquor company offered free on-site samples of the product coupled with extensive consumer advertising. These promotional efforts targeted to consumers created a demand for the product. Intermediaries such as wholesalers and retailers noticed consumer demand and consequently ordered the product to fulfill the requests of

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

consumers. This is an example of a pull strategy since the promotional efforts were directly aimed at end-users instead of intermediaries.

Pull/Push Combination Strategy

The type of strategy used depends on the product/market and product life cycle stage, and a variety of factors. For example, business-to-consumer companies tend to use pull strategies more than business-to-business companies (Kotler, 2003). Marketing strategies often rely on a combination of the two strategies in the promotion of a product. Few companies use a single strategy exclusively; however, one strategy may be emphasized over another. An example of a combined pull–push strategy can be seen in the pharmaceutical industry, where a push strategy is used by sending a sales force to promote drugs directly to physicians. This strategy is supplemented with a pull strategy consisting of consumer-targeted advertising in magazines and on television (*see* MARKET SEGMENTATION AND TARGETING). Although pharmaceutical companies typically emphasize push strategies, pull strategies are becoming more common in this industry (Parker and Pettijohn, 2005).

Bibliography

Kotler, P. (2003) *Marketing Management 11e*, Prentice Hall, Upper Saddle River.

Lamb, C., Hair, J. and McDaniel, C. (2006) *Marketing 8e*, Thomson Southwestern, Mason.

Parker, S.R. and Pettijohn, C.E. (2005) Pharmaceutical drug marketing strategies and tactics: a comparative analysis of attitudes held by pharmaceutical representatives and physicians. *Health Marketing Quarterly*, **22**, 27–43.

sales force strategy

William L. Cron and David W. Cravens

Introduction

A major objective of the twenty-first century company and its sales force strategy is to deliver superior customer value. Sales organizations must design and implement sales force strategies that are consistent with the firm's business and marketing strategies and create a customer-focused sales organization. The competitive and rapidly changing global business environment places top priority on aligning an organization's strategies to satisfy the escalating requirements of customers and to increase business performance.

Our objective is to examine the nature and scope of sales force strategy in the creation and delivery of superior value to priority customers. We begin by defining sales force strategy and examining the hierarchical nature of business strategy. Next, we consider customer relationships including customer prioritization and generic types of customer relationships. Attention is then directed to sales force capabilities and their role in creating a competitive advantage. Finally, we discuss the management support system and sales force structure essential in supporting and enhancing sales force capabilities.

Sales Force Strategy in Perspective

In this section we define sales force strategy and discuss the hierarchical nature of business strategy.

Sales force strategy. Sales force strategy is defined here as *the set of strategic decisions that determine as the areas on which the sales force will focus its attention and the role of the sales force in creating customer value that is consistent with the overall strategy of the firm and/or business unit*. Several aspects of this definition should be noted. First, this definition of sales force strategy indicates that the focus of the sales force should be on creating high levels of customer value (*see* CUSTOMER ANALYSIS) that enhances customer satisfaction and contributes to high levels of profitability. This focus marries the goals of the sales force with that of the overall organization in a market-driven organization. Second, the definition notes the hierarchical nature of sales force strategy, that is, a business strategy should drive a firm's MARKETING STRATEGY, which should guide the development of a firm's sales force strategy. Sales force strategy is conceived and executed within the context of an overall business strategy. This is to say that while a sales force's competencies and resources are considered when developing a firm's overall strategy, the sales force strategy in its totality is developed to best implement a firm's overall strategy. Third, sales force strategy is defined by a set of strategic decisions: customer prioritization, scope of customer relationship types, sales force capabilities portfolio, and a sales force structure and management support system to support these capabilities. The following sections further develop each of these sales force strategy elements.

Hierarchical context. As was mentioned earlier, the organizational context must be considered when examining sales force strategy. This includes considering the overall culture and basic philosophy of the organization. One aspect of culture that is of particular importance is the organization's basic orientation toward its markets. Sometimes referred to as a *market-driven strategy* (Cravens and Piercy, 2009), important characteristics include a firm's market orientation, customer focus, competitor focus, competitor intelligence, cross-functional coordination, and performance metrics (Cravens and Piercy, 2009).

Becoming market oriented involves getting the entire firm committed to the customer. The process requires the commitment and involvement of everyone in the company. The market-driven process begins with gaining an understanding of markets and customers by obtaining information on customers, competitors, and markets; analyzing the information through cross-functional involvement, deciding the strategic initiatives to pursue, and taking actions to deliver superior value to customers (Slater and Narver, 1994).

Achieving a customer focus is a demanding challenge. The organization must gain an understanding of customers' preferences and product

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

use requirements coupled with using the capabilities and resources of everyone in the organization to meet buyers' value requirements.

Competitor intelligence plays a pivotal role in market-driven strategy (*see* COMPETITIVE ANALYSIS). The process involves far more than marketing research studies. Relevant competitors and technologies must be determined, their short-term strengths and weaknesses and long-term capabilities identified, and strategies for delivering superior customer value formulated and implemented (Slater and Narver, 1994). Too often, management defines the competitive arena on far too narrow a basis. Consider, for example, the negative impact of digital photography on Polaroid and Kodak.

Cross-functional coordination and collaboration are essential in delivering superior customer value. Many traditional multilevel and functionally specialized organizations experience major problems in achieving cross-functional cooperation. Marketing and sales are illustrative of the difficulty of breaking down functional walls. Importantly, cross-functional teamwork is essential in the customer value delivery process (*see* CROSS-FUNCTIONAL TEAM).

There is a positive relationship between market orientation and business performance. Research studies that were conducted in the United States and several other countries show very encouraging performance findings. Thus, the return from implementing market-driven strategy in various firms appears to be worth pursuing.

Another contextual aspect of the sales force strategy is the hierarchical nature of an organization's goals, objectives, and strategies. As shown in Figure 1, there are interactions among the organization levels including corporate, business unit, marketing, and sales force strategy. While our interest is on sales force strategy, it is relevant to recognize these strategy level relationships.

Corporate strategy. The organization's strategy includes determining the range and purpose of the firm, its objectives, and the strategy actions and resources required to accomplish the objectives. As shown in Figure 1, the relationships among the levels should be recognized. Corporate strategy should determine the strategic direction of a firm, deploy resources, and set limits on what opportunities the firm will pursue whereas business unit and marketing strategies provide important feedback on market knowledge (*see* MARKET/INDUSTRY STRUCTURE), opportunities, and threats. An essential part of corporate strategy is determining the corporate vision in terms of what the business is all about (Cravens and Piercy, 2009).

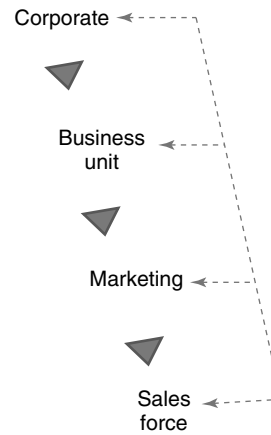


Figure 1 Business strategy levels and relationships.

Business unit strategy. On the basis of the composition of the corporation in terms of markets and products (goods and services), many companies are organized into specific units called *business units* or *strategic business units* (Cravens and Piercy, 2009). Each unit will focus on a product area and satisfies a set of similar customer group requirements. The business unit performs all (or most) of the basic business functions. This organizational unit has an identifiable management team and is responsible for sales and financial performance. Typically, there is close coordination between the corporate and business unit levels.

Marketing strategy. The design and implementation of marketing strategy requires detailed knowledge of corporate and business unit strategy to assure that marketing strategy is consistent with corporate and business unit strategies (*see* MARKETING STRATEGY MODELS). The chief marketing executive is likely to be involved in higher level strategy formulation. It may be recalled that the management guru,

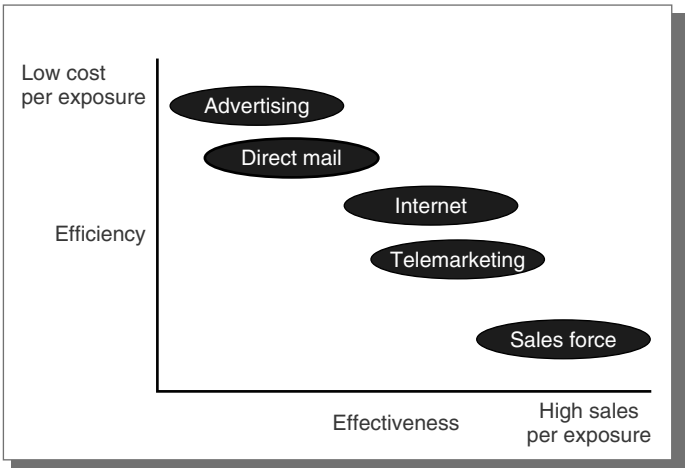


Figure 2 Comparing Go-To-Market alternatives.

Peter F. Drucker, defined marketing as a central dimension of the entire business.

Marketing strategy is concerned with developing a market vision for the organization, choosing market targets, setting objectives, and designing marketing program positioning strategies to respond to the value requirements of the customers in each target market (Cravens and Piercy, 2009).

For the sales force, a particularly important element of a firm's marketing strategy is its go-to-market strategy. A firm's go-to-market strategy defines the selling channel or combination of channels that can best execute a firm's marketing strategy (*see* GO-TO-MARKET STRATEGY). A wide array of selling channels exists today. In the past, firms typically chose a single channel to market, either their own direct sales organization or a group of distributors. Today, selecting the right selling channels is much more complicated. Companies reach out to their customers not only through direct sales and distributors but also through telemarketing, brokers, and manufacturer representatives, licensing arrangements, joint distributor agreements, dealers, new retail formats, and the Internet to name a few. Another factor that complicates the issue is the fact that companies sell through multiple channels to reach distinct customer segments and even to reach and service the same customers.

In general, companies are attempting to identify the best selling channels for executing their marketing strategy effectively and in the most cost-efficient way. Figure 2 compares several selling channels in terms of their effectiveness and efficiency. Note that a dedicated sales force is generally the least efficient (i.e., high cost per exposure), but the most effective (i.e., highest sales per exposure) means by which to connect to customers.

Sales force strategy. Marketing and sales force strategies require close coordination and collaboration. While sales force strategy is established within the context of a firm's marketing strategy, in the twenty-first century, the sales force plays an active role in planning many firm's marketing strategy, which requires close marketing and sales relationships.

Often, marketing and sales are organized as separate functions. Conflict may occur rather than collaboration between the functions, as indicated by many executive and academic thought leaders. This conflict is partially due to differences between the two functions in terms of what is considered critical and how a task is understood. For instance, marketing tends to be more product and long-term oriented, while sales is more customer and short-term oriented and possesses more granular market knowledge. Research has found that cooperation between

marketing and sales is harmed to the degree that these departments differ along these orientation and knowledge dimensions. On the other hand, an organization's market performance is found to be enhanced when both a product and a customer orientation are represented in the organization (Homburg and Jensen, 2007). These results support the involvement of sales in planning marketing strategy and the need for cooperation between sales and marketing.

Customer Relationships

In this section we discuss the types of customer relationships that guide the process of customer value creation. This decision is related to and a function of how a firm prioritizes its customers, that is, which customers receive the bulk of the sales force's attention.

Prioritization of customers. This process involves segmentation and assigning priorities to customer segments. A firm's marketing strategy will include a statement as to the firm's target market based on its market segmentation approach. Market segmentation involves grouping customers so as to have (i) similar characteristics, (ii) similar needs, and (iii) a propensity to respond similarly to a marketing program. Target marketing involves prioritizing the segments on which the firm will focus its planning and concentrate its marketing resources (*see* MARKET SEGMENTATION AND TARGETING).

The sales force strategy further refines a firm's customer focus by developing even more precisely defined customer segments, sometimes referred to as *microsegments*, and driving the prioritization process down to the individual customer level. At this level, in prioritization of customers individual customer factors such as a customer's present sales volume, future sales and profit potential, strategy, values, depth of present relationship, and organization may be considered. The purpose of such precise customer prioritization is to maximize efficiency in allocating sales force resources by maximizing CUSTOMER SOLUTIONS. Customer prioritization guides the next decision, namely, which is the type customer relationship that a selling firm wishes to establish with its target customers? This is discussed next.

Types of customer relationships. Customer relationships based on investments made by the supplier and investments by the buyer range from *transactional*, *consultative*, and *enterprise* type buyer/seller relationships (Cron and DeCarlo, 2009). The selection of relationship strategy guides initiatives for acquiring, maintaining, and developing customers. The decision determines which customers can be profitability served since the level of seller investment increases in the span from transactional to enterprise type relationships. Large companies may establish two or even three types of relationships with different customer groups.

Transactional relationship. This relationship consists of a buyer's need for a product with acceptable characteristics and quality, competitively priced, and a buyer/seller purchasing process convenient for the buyer. A favorable transactional relationship typically involves a personal relationship between both parties when the purchasing process is ongoing.

A recent study concluded that 68% of all firms focus on a transactional relationship with at least some of their customers. The overarching reason for emphasizing transactional relationships in these situations is because the buyers are already quite knowledgeable about the product and the importance of the product itself in meeting the buyer's objectives.

Consultative relationship. This type of relationship calls for a more collaborative approach between the salesperson and the buyer. The customer wants and needs sales and support effort which delivers value in addition to the actual product. Sometimes referred to as *solutions selling*, customers generally engage in consultative relationships to solve a problem for which they are not certain of the solution. The challenge is convincing the buyer of the value of the consultative relationship. Customer value is created by the following methods:

- helping customers understand their problems and opportunities in a new or different way.
- helping customers develop better solutions to their problems than they would have discovered on their own.

- acting as the customer's advocate inside the supplier's organization, ensuring the timely allocation of resources to deliver customized or unique solutions to meet the customer's special needs (Cron and DeCarlo, 2009, 56).

The time demands for the salesperson are much greater for consultative compared to transactional relationships due to analysis of customer needs/requirements and coordinating assembly of the internal resources necessary to deliver customer value. Time investment in the relationship is substantial for both the seller and the buyer.

One or more of the following factors point to the value of a consultative relationship:

- the product or service can be differentiated from competitive alternatives;
- the product or service can be adapted or customized to the needs of the customer;
- the customer is not completely clear about how the product or service provides solutions or adds value;
- the delivery, installation, or use of the product or service requires coordinated support from the selling organization;
- the benefits of the product or service justify the relatively high cost of consultative relationships (Cron and DeCarlo, 2009, 56–57).

These conditions signal an opportunity to utilize consultative selling in creating superior customer value.

Enterprise relationship. A key driver of the use of enterprise relationships is the major trend of customer firms to significantly reduce the number of suppliers and develop enterprise type relationships with the remaining key suppliers. This relationship:

> . . . is one in which the primary function is to leverage any and all corporate assets of the supplier in order to contribute to the customer's strategic success. In such a situation, both the product and the sales force are secondary, and the customer must be of strategic importance to the selling organization. (Cron and DeCarlo, 2009, 57).

Enterprise relationships may be designated by the supplier as key, strategic, major, national, or global accounts. The intent of the relationship is positive value generation for both the customer and the key supplier. Interestingly, the customer is frequently the initiator of an enterprise type of relationship.

Enterprise relationships generally involve supplier teams drawn from different functional areas and organizational levels. The team may be managed by an executive. Supplier/customer trust is important because of joint decision making, exchange of proprietary information, and other collaborative interactions. "As the buyer-seller relationship becomes more sophisticated and complex, the sales force's role as the primary point of contact between customer and supplier often diminishes" (Cron and DeCarlo, 2009, 59). Many of the premier industrial firms in the United States such as GE, IBM, Du Pont, and Monsanto have established enterprise relationships with customers such as American Airlines, Ford, Miliken, Procter & Gamble, and the federal government.

Having decided on a customer relationship mix for its target markets, the next issue to be addressed is developing the sales force capabilities to support a selected type of customer relationship. Some fundamental sales force capabilities are described in the next section.

Sales Force Capabilities

Having laid the groundwork by identifying the types of customer relationships the sales force is tasked with establishing, management is now in a position to address a core issue of sales force strategy, which is identifying the set of sales force capabilities necessary to successfully implement our customer relationship strategy. A capability refers to a firm's capacity to deploy and leverage resources to create customer value and a competitive advantage. Just as an individual may have the capability to play a violin or basketball, or speak German, so a sales force may be capable of landing new customers, sustaining and growing key customer relationships, and building deep trusting relationships.

There are two key aspects of capabilities to which attention should be devoted: resources and coordination/deployment of resources. Resources are the productive assets available to a firm (*see* MARKET-BASED ASSETS). There

are three principal types of resources: tangible (e.g., number of sales forces and sales offices), intangible (e.g., brand awareness and company equity), and human resources (e.g., sales force knowledge and attitude). It is important to recognize, however, that resources by themselves do not result in competitive advantage. Rather, resources must be leveraged to create a capability. It is capabilities that are the essence of superior performance.

Capabilities require the expertise of people to integrate with resources to establish routines, which are predictable patterns of activity, typically involving coordinated action by individuals (Barney, 1997). The focus of these routines is on creating value. So, an important capability for Honda Motor Company is its expertise in developing and manufacturing engines. This expertise has taken the firm into both automobiles and motorcycles, has been valuable to customers, and has created a competitive advantage for Honda.

Capabilities are hierarchical in nature. That is, broad organizational capabilities such as Honda's engine capability can be disaggregated into more specialist capabilities. Thus, we can think of the sales force as having certain capabilities such as its capability to acquire new customers, develop deep customer relationships, gain customer trust, and retain valuable customers.

Considered from this perspective, the sales force can be thought of as having routines for performing a set of important capabilities. Deciding on the nature and scope of these routines then is critical to developing sales force strategy. It is important to note, however, that the nature of these capabilities will be closely associated with the types of relationships the seller has with its customers. Table 1 describes how four sales force capabilities may differ depending on the nature of the seller–customer relationship. Note that trust is built in different ways in a transactional relationship, for instance, than it is in a consultative relationship. This is one of the reasons why companies may find it difficult to change from a transactionally oriented sales force to one that is consultative. Not only do new routines need to be learned and developed, but old ones need to be undone.

How successful a sales force will be in developing capabilities that will lead to superior customer value and competitive advantage will depend on the organizational structure in which sales force capabilities are employed and management systems for supporting and enhancing the continuous development of these capabilities. This is where we direct our attention in the next section.

Sales Force Structure

Organizational structure should support the appropriate sales force activities associated with the sales force competencies. In other words, the structures should create stability and continuity of sales activities to support the development of routines. Importantly, the organizational design needs to facilitate coordination of sales force activities allocated to different members of the organization and different departments.

In this section, we discuss horizontal and vertical structures, team selling, and sales channel strategy.

Horizontal structure. A company may elect to outsource the sales function or use a company sales force. Alternative sales force structures include organizing by geography, type of product, type of customer, or selling function organizations (Cron and DeCarlo, 2009).

Outsourcing. Management may decide to utilize an independent organization to perform the sales function. The use of a manufacturer's representative is a popular option (Anderson and Trinkle, 2005). Representatives are independent businesses given exclusive contracts to perform the selling function within a specified geographic area. They take neither ownership nor physical possession of the products they sell and are always compensated by a commission on sales volume. The financial advantage of manufacturer representatives is in reduced capital requirements and favorable cash flow in that they are compensated only when sales are obtained.

Other outsourcing options involve contracting with value chain members to perform the sales function. For example, a producer may arrange for a distributor to provide sales coverage to retailers. The two key limitations of all outsourcing options is limited control by the firm utilizing the outsourcing option and higher

Table 1 Sales force capabilities and customer relationship type.

Sales Force Capabilities	*Transactional Relationship*	*Consultative Relationship*	*Enterprise Relationship*
New customer acquisition	Prospecting is broadly targeted to general business categories. Appeal focuses on the product and/or price.	Prospecting is more targeted with referrals playing a larger role. The focus is on customer problems and solutions typically with some degree of customization of the seller's offering.	Prospecting is about identifying opportunities to collaborate with existing, strategic customers on mission-critical issues. The focus is on the opportunity to bring to bear supplier resources and capabilities on the customer's strategic issues and opportunities.
Customer relationship development	Building personal relationships with key customer personnel and growing the breadth of product and product lines that customers purchase.	Based on implementation of the solution and providing ongoing support and monitoring of evolving customer requirements.	Supplier leads deployment of the solution with customer oversight and is often responsible for running system following deployment and addressing any contingencies that might arise.
Customer trust building	The offering performs as promised. A deeper personal and professional relationship is developed between key participants and any problems are quickly and adequately resolved.	Sufficient enterprise-level value creation occurs based on cross-enterprise, cross-functional capabilities and commitments between supplier and customer.	The selling organization has the capability to provide insights into the customer's execution gaps, develop a joint vision of the relationship, and is willing to invest the resources necessary to execute and sustain the visions.
Customer retention	How well the product performs given its price and competitive offerings and the ability of the sales force to maintain key personal relationships.	The quality with which the selling organization provides support following deployment of the offering, which may include a tracking system of performance, sharing new insights, and providing cutting-edge upgrades.	Willingness to make changes to create value, sufficient potential to justify investment, cross-functional capability and commitment, and capability of organizational values.

Table 2 Sales force structure specialization.

Organizational Structure	*Advantages*	*Disadvantages*
Geographic	Low cost No geographic overlap No customer overlap	Limited product line knowledge Limited customer knowledge Lack of management control over product or customer emphasis
Product	Product knowledge Control over product emphasis	Low geographic efficiency Customer duplication Geographic duplication
Customer	Deeper customer knowledge Control over customer emphasis	High cost Less product knowledge More geographic duplication Difficult coordination with product managers
Functional	Effectiveness in performing selling activities	Coordination Geographic duplication Customer duplication

total costs compared to a dedicated sales force when sales exceed a certain threshold.

Specialization. A dedicated sales force may be specialized in a variety of ways. In addition to concern with the effectiveness with which sales capabilities are developed and supported, adaptability and efficiency are also important concerns when designing sales force structure. Sales organization adaptability is concerned with how quickly a company can react to product and market changes. Efficiency reflects the ease and speed with which key sales activities, for example, sales calls, proposals, product customization, can be performed.

Common structures for a specialized sales force include geographic, type of product, type of customer, or selling function organizations. Each type of specialization helps support critical sales force capabilities so as to meet the needs of customers and to excel against direct competitors. While each type of sales force specialization has its advantages, there are also disadvantages to each type of organization which management will attempt to minimize. The advantages and disadvantages of each type of sales force organization are listed in Table 2. There has been a trend in recent years, however, toward customer specialization and away from a purely geographic organization.

Team selling. The use of sales teams requires attention to the design of the team selling structure (Cron and DeCarlo, 2009). Teams may be used with large and more attractive customers owing to the financial cost associated with team selling. They are normally used with enterprise relationships and may be appropriate in certain consultative relationships. A team may comprise people from several functions and a team leader. These designs include challenges in coordination and collaboration but may offer sufficient advantages to overcome the limitations.

Vertical structure. The vertical structure of the sales organization consists of the number of levels of management between the chief sales executive and salespeople. The greater the number of levels, the fewer people each manager supervises. Multiple management levels were considered essential in the past to facilitate communications and interactions up and down the organization. Significant advances in information technology have substantially improved communications between and across levels. Nonetheless, in sales the range of span of control is quite large across companies with as few as 3 salespeople per manager to as many as 30. The most common range is probably between 7 and 12.

The issues of number of management levels and span of control are important because they directly affect sales force costs and revenues.

Determining the number of management levels and number of people supervised may depend on relevant situation-specific factors (e.g., product support requirements and type of customer relationship). In general, a flat organization and a wider span of control are appropriate when the managerial task is less time consuming. Factors that increase the amount of time that a manager needs to spend with individual salespeople include:

1. wide geographic dispersion of salespeople;
2. many inexperienced salespeople; and
3. extensive coordination required across salespeople or with others in the organization (Zoltners, Sinha, and Lorimer, 2004, 168).

Sales territory design. An important aspect of designing an effective sales organization is the design of sales territories. Customers have a level of activity and coverage that is needed to properly service them, while salespeople have an activity capacity. Sales territories consist of the customers located in a specified geographic area that are assigned to an individual salesperson. While territories are often referenced in geographic terms, the defining element of a territory is the set of customers in the geographical area.

A properly aligned sales territory is one in which customers receive the appropriate level of sales force attention and the workload is balanced across salespeople. There appears to be a great opportunity for improvement with respect to proper territory alignment. A study of over 4800 sales territories found that well over half of the territories were not the right size. Approximately 25% of the territories were too large to be effectively covered, while 31% were too small to keep a salesperson sufficiently busy with productive work (Zoltners, Sinha, and Lorimer, 2004, 274).

Sales territory alignment is an ongoing course of business requiring frequent minor adjustments, but major realignments are also needed every few years at many companies. Conditions that create major territory realignments include when a company changes its type of customer relationship focus, significantly increases/decreases the size of its sales force, creating new products, and there are significant mergers and acquisitions. Experts suggest that companies audit their territory alignments annually.

Management Support System

The sales organization's management support system plays a vital role in the performance of the organization through its impact on sales force capabilities. In this section, we discuss management control processes, metrics alignment, the motivation system, and talent development.

Management control processes. Behavior-based and outcome-based sales force management control systems were proposed over 20 years ago (Anderson and Oliver, 1987). Behavior-based management control consists of specific sales management activities (monitoring, directing, evaluating, and rewarding salespeople), and the extent to which managers perform these activities. Compensation consists of relatively high percentages of fixed salaries for salespeople. Outcome-based control utilizes relatively high levels of incentive compensation for salespeople and relatively limited sales manager control activities.

Behavior control is focused on salesperson selling behavior whereas outcome control is more concerned with salesperson results. Under outcome control, salespeople have very limited interaction with sales managers during their sales processes. Minimal monitoring and manager direction occur. Behavior control is primarily concerned with the processes salespeople use to obtain sales results.

There is substantial research support for a positive impact of behavior control on salesperson performance, which in turn displays a positive relationship with sales unit effectiveness. Salesperson performance comprises behavior and outcome performance. Behavior performance has been found to have a strong positive relationship with salesperson outcome performance, which has a positive impact on sales unit effectiveness.

The intent is not to suggest that behavior control is a favorable management process and outcome control has a negative impact on salesperson performance and sales unit effectiveness. Instead, behavior control appears more

appropriate for nontransaction type buyer/seller relationships.

Metrics alignment. An important aspect of sales force strategy is tracking how well the strategy is performing. Marketing and sales force metrics are of strong interest in many companies driven by concerns as to the measurement of performance and relating it to business performance. Metrics has been assigned a top priority research issue by the Marketing Science Institute (MSI) (*see* MARKETING METRICS). MSI is a not-for-profit research organization funded by corporate members who set research priorities. A key issue is developing measures for marketing and sales that are related to business performance.

Illustrative metrics include tracking position relative to competition, CUSTOMER LIFETIME VALUE (CLV), strategy effectiveness, and financial performance. Normally, multiple measures are recommended. Sales organization metrics should be compatible to those used by the business and the marketing function. Sales metrics are needed but should be extended to include areas such as account profitability and contribution to profit.

Sales managers need to develop knowledge as to what strategy actions lead to improved sales productivity. Sales management behavior control is an example since research findings indicate a positive relationship with sales organization effectiveness. Effectiveness has been measured based on sales, market share, profitability, and customer satisfaction.

Motivation system. Motivating salespeople is a continuing challenge for sales organizations. While motivation is in part self-generated, development of an internal organizational motivation system is an essential aspect of sales force strategy. We know that motivation is a very relevant influence on job performance (*see* MOTIVATION RESEARCH). The starting point is gaining an understanding of the factors that influence a person's motivation.

Accumulated knowledge on motivation points to the interplay between three factors affecting the amount of effort someone will put into an activity: (i) the relationship between effort and performance, referred to as *expectancy*; (ii) the relationship between performance and rewards, called *instrumentality*; and (3) the importance of receiving more of a certain reward, known as *valence* (Cron and DeCarlo, 2009). Combined, these three factors provide a framework for managers to understand the internal process by which people are motivated to put forth extra effort.

There are several sales management actions that may have a positive impact on motivation of sales personnel. These actions include helping salespeople set challenging yet achievable goals; rehearsing desired behaviors through one-on-one coaching, setting fair and challenging quotas, incentive programs, and recognition programs (Cron and DeCarlo, 2009, 249–251).

Talent development. It is generally agreed that talent development is the most important dimension of sales management. Talent development focuses on the issues of having the right people with the right values, skills, and capabilities. It encompasses recruiting, training, coaching, and culture development. Underlying successful execution of each sales force capability is the appropriate level of sales force knowledge, attitude, and values.

Hiring is a critical activity because it affects company results, management effectiveness, and company culture. A report by the Harvard Business School estimates that a bad hiring decision can cost a company three times the salesperson's annual compensation, including expenses, training costs, benefits, and incentive pay. Likewise, new recruits bring their own values to the company. Successful people can become heroes for the company and successful "war stories" become the legends of the company.

Companies can successfully take different approaches to talent development. Selling organizations such as Procter & Gamble, Johnson & Johnson, and Xerox primarily hire recent college graduates for the sales positions and train them thoroughly to be successful. Training programs are considered a corporate capability that generates a competitive advantage for these companies. In contrast, other companies hire only people with sales experience. They expect the new hires to integrate quickly and contribute to the selling effort. A fundamental question that

management should ask itself is which salesperson characteristics can be developed through training and which are innate? It should be borne in mind that only skills and knowledge can be trained.

Conclusion

We have defined and examined the various dimensions of sales force strategy, which consists of a set of strategic decisions that determine who the sales force will sell to and the role of the sales force in creating customer value that is consistent with the overall strategy of the firm and/or business unit. At the center of the strategy process is the delivery of superior customer value to targeted customers. The past role of the sales force has changed from primarily implementing strategy to becoming an active participant in marketing strategy development.

Sales force strategy is linked to corporate business units, and marketing strategies in an interrelated hierarchy of relationships. Marketing and sales force strategies require close coordination and collaboration. Prioritization of customers is a key joint effort which is intended to result in determining the types of customer relationships pursued by the firm. The options include *transactional*, *consultative*, and *enterprise* relationships. The determination of the relationship strategy (or strategy combination) provides the guidelines for actions in acquiring, maintaining, and developing customers.

Sales force strategy requires developing sales force capabilities to support each type of strategic relationship. These include new customer acquisitions, customer relationship development, customer trust building, and customer retention. Sales force capabilities may vary depending on the nature and requirements of the seller–customer relationship. These demands make it difficult for a company to change from a transactional oriented sales force to one that is consultative.

Developing the necessary capabilities to achieve superior customer value and competitive advantage is very dependent on selecting the appropriate organizational structure in which sales force capabilities are employed and management systems for supporting and enhancing the continuous development of these capabilities. Horizontal structure choices include possible outsourcing, extent of specialization (geographic, type of product, type of customer, and/or selling function), and use of team selling. Vertical structure decisions need to address the number of levels and span of control of sales management and sales territory design.

The management support system consists of choices concerning control processes, metrics alignment, the motivation system, and talent development. Control focuses on the extent of behavior versus outcome control. Metrics choices are essential in tracking how well the strategy is performing. The motivation system determines which management actions will be used to create a positive impact on motivation of sales personnel. Talent development focuses on the issues and choices of having the right people with the right values, skills, and capabilities as members of the sales force.

A successful sales force strategy requires appropriate choices concerning the various customer relationship, sales force capabilities, organizational structure, and management support system decisions. These strategy dimensions are closely related and require careful integration. The end result should provide a successful superior customer delivery system.

Bibliography

Anderson, E. and Oliver, R. (1987) Perspectives on behavior-based versus outcome-based salesforce control systems. *Journal of Marketing*, **51**, 76–88.

Anderson, E. and Trinkle, R. (2005) *Outsourcing the Sales Function*, Thomson/South-Western, Mason.

Barney, J. (1997) *Graining and Sustaining Competitive Advantage*, Prentice Hall, Englewood Cliffs.

Cravens, D.W. and Piercy, N.F. (2009) *Strategic Marketing*, 9th edn, McGraw-Hill/Irwin, Burr Ridge.

Cron, W.L. and DeCarlo, T. (2009) *Sales Management*, 10th edn, John Wiley & Sons, Inc, New York.

Homburg, C. and Jensen, O. (2007) The thought worlds of marketing and sales: which differences make a differences. *Journal of Marketing*, **71**, 124–142.

Slater, S.F. and Narver, J.C. (1994) Market orientation, customer value, and superior performance. *Business Horizons*, **37**, 22–27.

Zoltners, A., Sinha, P., and Lorimer, S. (2004) *Sales Force Design for Strategic Advantage*, Palgrave Macmillan, New York.

services marketing strategy

Valarie A. Zeithaml, Mary Jo Bitner, and Dwayne D. Gremler

Services marketing strategy focuses on delivering processes, experiences, and intangibles to customers rather than physical goods and transactions. It involves integrating a focus on the customer throughout the firm and across all functions. All company functions – marketing, selling, human resources, operations, and R&D – must work together to create effective services marketing strategy. Rather than the traditional goods marketing focus on transactions and exchange, services marketing strategy is centered on the customer, usage, and relationships (Vargo and Lusch, 2004a).

Services, which can be defined as *deeds, processes, and performances*, fall into several categories. Many services, such as hotels, transportation, and health care, are offerings in and of themselves and are the primary revenue-producing activities of the firms. Another category of service is customer service, which includes the service provided in support of a company's core products. Typically, customer service does not directly produce revenue but rather addresses customer requests, questions, and complaints, besides providing answers and solutions. Service can also be a value-add for manufactured products – many companies provide training, installation, and repair services for the goods they produce – often for a fee. Finally, many services are derived from or are provided by manufactured products such as cell phones, computers, software, and mobile phones. In early writings on services, scholars distinguished services from goods by noting that they were intangible, perishable, variable, and that the producer and consumer were inseparable. Recently, it has been suggested that these distinctive characteristics should not be viewed as unique to services but that they are also relevant to goods, that "all products are services," and that "economic exchange is fundamentally about service provision" (Vargo and Lusch, 2004b). Although this view is rather abstract, it does suggest that all types of organizations can gain valuable insights from services marketing frameworks, tools, and strategies.

The four Ps of marketing (product, price, promotion, and place) are only partially adequate to conduct effective services marketing strategy. Three additional Ps – *people, process*, and *physical evidence* – are also needed. Because services are usually produced and consumed simultaneously, customers are often present in the firm's factory, interact directly with the firm's personnel, and are actually part of the service production process. Therefore, all human actors (the P corresponding to people) play a part in service delivery and thus influence the customer's perceptions. The firm's personnel, the customer, and other customers in the service environment each provide cues to the customer regarding the nature of the service itself. Employee dress, personal appearance, attitudes, and behaviors all influence the customer perceptions of the service. Physical evidence pertains to the environment in which the service is delivered; all tangible components that facilitate performance or communication of the service also affect services marketing. This includes all the tangible representations of the service such as brochures, letterheads, business cards, report formats, signage, and equipment, and the *servicescape*, the physical facility where the service is offered. This P (physical evidence) is needed for services because customers often have little on which to judge the actual quality of an intangible offering and thus will rely on any tangible components of the service offering. Finally, process – including the operating systems, procedures, mechanisms, and flow of activities by which the service is delivered – is an element of the services marketing mix. This P (process) addresses how the service is delivered, which, in many cases, may be perceived by customers to be as important as the outcome of the service. In the discussion that follows, the impact of these additional three Ps in services marketing strategy will be obvious.

Gaps Model of Service Quality

Executives of services organizations have long struggled with how to approach service design and delivery in an organized manner. The dominant approach to viewing the delivery of service quality in a structured and integrated way is called the *gaps model of service quality*

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

(Parasuraman, Zeithaml, and Berry, 1985). The gaps model positions the key concepts, strategies, and decisions in delivering quality service in a manner that begins with the customer and builds the organization's tasks around what is needed to close the gap between customer expectations and perceptions. The gaps model provides a comprehensive and integrating framework for delivering service excellence and customer-driven service innovation.

The model is particularly relevant in service strategy because it captures the cross-functionality inherent in service management. Although the authors are marketing academics and their original publications appeared in marketing journals, their work has been widely cited and used across academic disciplines and implemented in different functions within organizations. The model draws heavily from logic, theories, and strategies in operations, human resources, marketing, and increasingly from information systems.

The model, illustrated in Figure 1, is anchored on the customer and integrates customer focus throughout all gaps within the model. As depicted in the model, a firm's primary goal should be to meet or exceed customer expectations, and strategies used to achieve that objective (whether operations, human resource, or technology-based) are all focused on the customer. Every gap and every strategy used to close the gaps in the model retains a focus on the customer at its core.

The central focus of the model is the *customer gap* – the difference between customer expectations of what will be delivered and perceptions of the service as it is actually delivered. The other four gaps in the model are known as the *provider gaps* and each represents a potential cause behind a firm's failure to meet customer expectations. Closing the customer gap – delivering quality service – at the top of the model is a complex undertaking involving many different organizational and employee skills and tasks. These tasks can be sorted into four other gaps – the provider gaps – each of which needs to be closed in order to close the customer gap.

The following four provider gaps, shown under the horizontal line in Figure 1, are the underlying causes behind the customer gap:

Gap 1: The listening gap
Gap 2: The service design and standards gap
Gap 3: The performance gap
Gap 4: The communication gap.

At its most basic level, the logic of the model suggests that the customer gap is a function of any one or all of the four provider gaps. Early publications of the gaps model enumerated the complex reasons that cause each of these provider gaps. Later publications and a leading services marketing textbook (Zeithaml, Bitner, and Gremler, 2009) have further elaborated on the gaps by delineating specific strategies for closing each of them. We will expand briefly on key strategies used to close each of the gaps.

The Customer Gap

The customer gap is the heart of the gaps model. It represents the difference between customer expectations and perceptions of service performance. The model suggests that closing this gap by matching or exceeding customer expectations will result in the achievement of service quality from the customer's perspective. In the years since the introduction of the model, there has been significant focus on both customer expectations and perceptions in terms of conceptualizing these constructs, developing measures for them, and studying their effects.

A prominent focus of both academic and practical research has involved identifying the dimensions of service quality. Considerable exploratory and empirical work resulted in the identification of five dimensions:

1. *Assurance*: Knowledge and courtesy of employees and their ability to inspire trust and confidence.
2. *Empathy*: Caring, individualized attention the firm provides its customers.
3. *Reliability*: Ability to perform the promised service dependably and accurately.
4. *Responsiveness*: Willingness to help customers and provide prompt service.
5. *Tangibles*: Appearance of physical facilities, equipment, personnel, and communication materials.

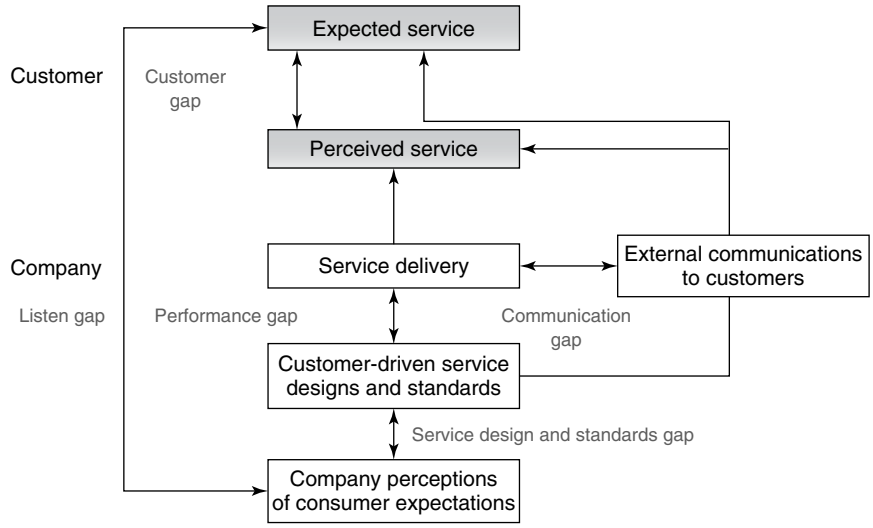


Figure 1 Gaps model of service quality.

The five dimensions of service quality have been captured in a questionnaire called *SERVQUAL*, consisting of a total of 21 items measuring these dimensions. The SERVQUAL measure has been applied in and adapted to many industry settings. Related streams of research have developed in parallel to study service encounters, customer satisfaction, customer loyalty, and their relationships with service quality. None of these now-prominent streams of research existed prior to the 1980s, and all continue to spawn research today.

Gap 1 – The Listening Gap

Provider gap 1, *the listening gap*, is the difference between customer expectations of service and company understanding of those expectations. A primary cause in many firms for not meeting customers' expectations – that is, the customer gap – is that the firm lacks accurate understanding of exactly what those expectations are. Many reasons exist for managers not being aware of what customers expect: They may not interact directly with customers, they may be unwilling to ask about expectations, or they may be unprepared to address them. Closing the listening gap requires that management or empowered employees acquire accurate information about customers' expectations. Customer expectations must be assessed accurately before new services are developed, and they must be tracked after the services are introduced.

Figure 2 lists the three key strategies for closing the listening gap. Each of these strategies is backed by research and practical applications. The first strategy is to listen to customers in multiple ways through customer research and employee upward communication. Such research includes the full range of traditional marketing research methods such as surveys, focus groups, and complaint handling. There have also been research methods uniquely useful in service situations such as SERVQUAL surveys, mystery shopping, and critical incidents analysis. A distinguishing factor between marketing research on goods and services is that services research must capture human performance. Whereas goods research can evaluate goods independent of the individuals who create them, service is often created in the interaction between customers and contact personnel. The behavior of personnel can be highly variable across individuals as well as with employees from day to day, so constant monitoring must occur. For that reason, additional techniques are needed to assess and feedback information about the performance of individuals. *Mystery shopping* – hiring people

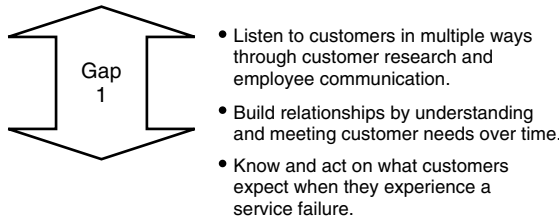


Figure 2 Strategies for closing the listening gap.

to pose as customers to evaluate performance – is typically used in restaurants and other retail service settings. *Critical incidents research*, in which a customer recalls and discusses both satisfying and unsatisfying experiences with a service provider and its employees, is particularly useful in examining and improving service encounters. Another marketing research approach that is particularly useful in service firms is the *trailer call*, a short survey that follows ("trails") a service event or encounter. The trailer call offers quick feedback on employees and also allows a company to fix its processes in a timely fashion.

The second strategy that closes provider gap 1 is to focus on building relationships by understanding and meeting customer needs over time. In firms where customers and companies have interpersonal contact, this can involve many different strategies: learning customers' names, their businesses, their industries, and their histories with the firm. Even in direct marketing or online situations, a firm can develop a virtual relationship with customers by learning their preferences and history. The stronger the firm's relationship with its customers, the better is the firm's ability to listen to customers (and thus close the listening gap).

The final key factor associated with provider gap 1 is lack of service recovery, or a failure to understand and act on what customers expect when there is a service failure. Even the best companies, with the best of intentions and clear understanding of their customers' expectations, sometimes fail. It is critical for an organization to understand the importance of service recovery – why people complain, what they expect when they complain, and how to develop effective service recovery strategies for dealing with inevitable service failures. Such strategies might involve a well-defined complaint-handling procedure and an emphasis on empowering employees to react on the spot, in real time, to fix the failure; at other times, it involves a service guarantee or ways to compensate the customer for the unfulfilled promise. Firms that learn from their failures – which often result from not fully understanding their customers' expectations – can reduce or eliminate the listening gap.

Gap 2 – The Design and Standards Gap

Closing gap 1 through research and effective management of customer relationships is necessary, but not sufficient, for achieving service excellence. Even when a company has a thorough and ongoing understanding of its customers' expectations, it is still very possible, in fact, quite easy, to fail to deliver quality service. Focusing on gap 2, the design and standards gap, is the next step toward ensuring against such failure. This gap is concerned with translating customer expectations into actual service designs and developing standards to measure service operations against customer expectations.

Figure 3 summarizes several key strategies for closing gap 2. The first strategy is to employ well-defined new-service development and innovation practices for designing services. Some have referred to this as *formalization* of a "services R&D" practice. While standardized new-product development processes and R&D are common in technology and manufacturing, they are still quite rare in services (for a major exception, we note the investment of the IBM Corporation in service innovation research through its global research labs). A formalized process typically involves a series of steps beginning with strategy formulation and idea

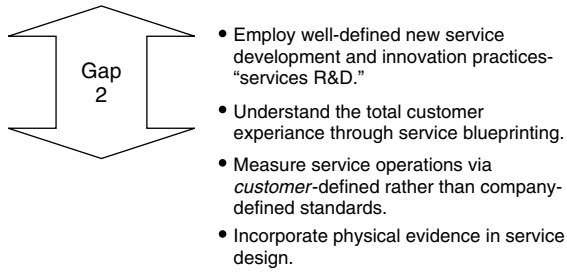


Figure 3 Strategies for closing the design and standards gap.

generation and ending with full-scale implementation. Because of the nature of services (their process orientation, intangibility, cocreation by customers), it is more challenging to engage in these typical steps that are so well established in other industries. However, it is clear that following a well-defined process, engaging customers along the way, and carefully planning and prototyping the complexities of service implementation are all essential in ensuring service designs that meet customer expectations. Building prototypes of services and planning for eventual full-scale implementation again means that operations, marketing, and, in many cases, human resource functions must work together.

A second strategy for closing gap 2 relates to understanding the total customer experience and designing all elements of that experience in ways that meet or exceed customer expectations. This involves considering everything that occurs from the moment the customer engages the service through the entire length of the service experience. Common elements of the service experience that need to be designed include customer-facing processes, the physical space where the service is delivered, and the interactions between service employees and customers. Viewing these operational elements from the customer's perspective and designing them to be consistent with expectations, or to reinforce a desired service image, are critical to closing gap 2. Because of the special challenges inherent in designing services, techniques such as *service blueprinting* have evolved to aid in the design process (Zeithaml, Bitner, and Gremler, 2009). The purpose of a service blueprint is to make a complex and intangible service concrete through its visual depiction of all of the steps, actors, processes, and physical evidence of the service. The key feature of service blueprints is their focus on the customer – the customer's experience is documented before any of the internal processes are determined.

A third strategy for closing gap 2 involves measuring service operations via *customer-defined standards*. These are standards set to correspond to customer expectations rather than to firm-focused goals. The quality of service delivered by customer-contact personnel is critically influenced by the standards against which they are evaluated and compensated. Standards signal to contact personnel what the management priorities are and which types of performance really count. When service standards are absent or when the standards in place do not reflect customers' expectations, quality of service as perceived by customers is likely to suffer. Customer-defined standards can either take the form of operational (hard) or perceptual (soft) standards.

The final strategy that closes gap 2 involves the use of physical evidence in service design and in meeting customer expectations. This includes everything tangible in the service-delivery process, such as business cards, reports, signage, Internet presence, equipment, and facilities used to deliver the service. The servicescape, the physical setting where the service is delivered, is a particular focus of physical evidence and is critical in industries such as restaurants and hotels to communicate about the service and make the entire experience pleasurable. In these cases, the servicescape plays a variety of roles, from serving as a visual metaphor for what the company stands for to actually facilitating the activities of both consumers and employees.

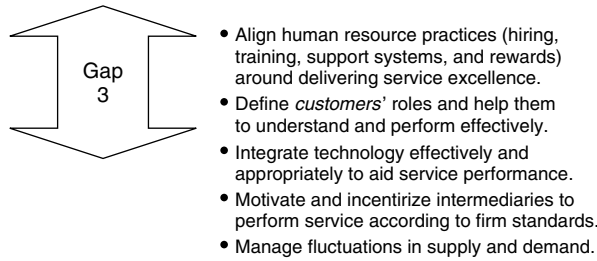


Figure 4 Strategies for closing the service performance gap.

Given the importance of physical evidence and its potentially powerful influence on both customers and employees, it is important for firms to think strategically about the design and management of the tangible evidence of service.

To accomplish the type of customer-focused service design described above means that marketing and operations functions within the firm must at least share information, and, in the best case, actually collaborate in designing the service and setting standards for its delivery. Because of the interpersonal nature of many services it is critical to design and plan for the human element as well; thus, effective service design and customer-defined standards will reflect collaboration with the human resource function in the organization. These types of cross-functional collaborations are not as essential in goods marketing where process, people, and physical evidence (the new-services marketing mix elements) are not as critical in defining the product.

Gap 3 – The Performance Gap

Although a company may have closed both the listening gap and the service design and standards gap, it may still fall short of providing service that meets customers' expectations if it is unable to deliver service in the way the service was designed. Gap 3, the service performance gap, must also be closed to make sure there is no discrepancy between customer-driven service design and standards and actual service delivery. Even when guidelines exist for performing service well and treating customers correctly, high-quality service performance is not a certainty. Standards must be backed by appropriate resources (people, systems, and technology) and also must be enforced to be effective – that is, employees must be measured and compensated on the basis of performance along those standards. Thus, even when standards accurately reflect customers' expectations, if the company fails to provide support for those standards – if it does not facilitate, encourage, and require their achievement – standards do no good. When the level of service delivery falls short of the standards, it falls short of what customers expect as well. Narrowing gap 3 by ensuring that all the resources needed to achieve the standards are in place reduces the customer gap.

The key strategies for closing gap 3 are depicted in Figure 4. The first strategy is to align the firm's human resource strategies around delivering service excellence. In particular, in order to deliver service as it was designed a firm needs to ensure that employees are willing and able to deliver quality services and that they are motivated to perform in customer-oriented, service-minded ways. In creating such a workforce, an organization must hire the right people, develop those people to deliver service quality, and retain the best people. To effectively deliver service quality, considerable attention should also be focused on recruiting and hiring the right service personnel. Service employees need two complementary capacities: *service competencies* – the skills and knowledge necessary to do the job – and *service inclination* – an interest in doing service-related work. Once the right people are in place, to provide quality service they need to be developed through ongoing training in the necessary technical skills and interactive skills. An organization that hires the right people and trains and develops them to deliver service quality must also work to retain them. If a

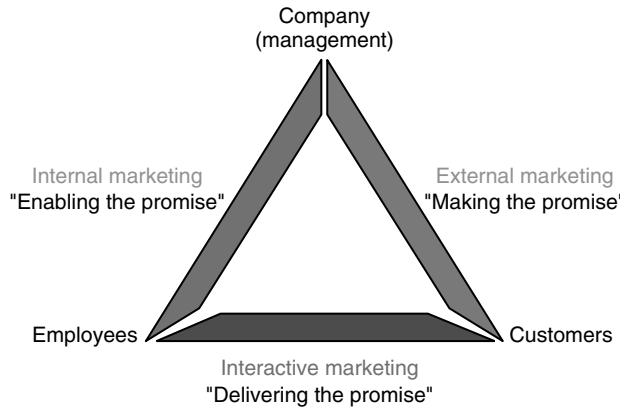


Figure 5 The services marketing triangle.

company wants the strongest service performers to stay with the organization, it must reward and promote them. Organizations use a variety of rewards to retain the best employees; traditional approaches such as higher pay, promotions, and one-time monetary awards or prizes are often linked to service performance.

Services marketing is about promises made and promises kept to customers. A strategic framework known as the *services triangle* (Figure 5) visually reinforces the importance of people in the ability of firms to keep their promises and succeed in building customer relationships (Bitner, 1995; Kotler, 1994; Grönroos, 2007).

The triangle shows the three interlinked groups that work together to develop, promote, and deliver services. These key players are labeled on the points of the triangle: the company (or SBU (small business unit) or department or "management"), the customers, and the providers. Providers can be the firm's employees, subcontractors, or outsourced entities who actually deliver the company's services. Between these three points on the triangle, three types of marketing must be successfully carried out for a service to succeed: external marketing, interactive marketing, and internal marketing.

On the right side of the triangle are the *external-marketing* efforts that the firm engages in to set up its customers' expectations and make promises to customers regarding what is to be delivered. Anything or anyone that communicates to the customer before service delivery can be viewed as part of this external-marketing function. But external marketing is just the beginning for services marketers: promises made must be kept. On the bottom of the triangle is what has been termed *interactive marketing* or *real-time marketing*. This is where promises are kept or broken by the firm's employees, subcontractors, or agents. People are critical at this juncture. If promises are not kept, customers become dissatisfied and eventually leave. The left side of the triangle suggests the critical role played by *internal marketing*. Management engages in these activities to aid the providers in their ability to deliver on the service promise: recruiting, training, motivating, rewarding, and providing equipment and technology. Unless service employees are able and willing to deliver on the promises made, the firm will not be successful, and the services triangle will collapse.

All three sides of the triangle are essential to complete the whole, and the sides of the triangle should be aligned – that is, what is promised through external marketing should be the same as what is delivered; and the enabling activities inside the organization should be aligned with what is expected of service providers.

For many services, customers are participants in service production and cocreators of value and, therefore, play a key role in the service-delivery process – that is, customers themselves can influence whether the service

meets customer-defined specifications and can potentially contribute to the widening of gap 3. Therefore, a second strategy for closing the performance gap is to define customers' roles and assist them in understanding and performing their roles effectively. Sometimes customers widen gap 3 because they lack understanding of their roles and exactly what they are to do in a given situation or because they are unwilling or unable to perform for some reason. To reduce this gap the organization needs to clearly define and communicate what the customer's role entails – in essence, the customer's job description. Once the customer's role is clearly defined, the firm needs to help facilitate that role. In a sense, the customer is a "partial employee" of the organization, and strategies for managing customer behavior in service production and delivery can mimic, to some degree, the efforts aimed at service employees discussed in the previous paragraph.

A third strategy for closing gap 3 involves integrating technology effectively and appropriately to aid service performance. For service workers (and customers) to be efficient and effective in performing their jobs, technology that facilitates their efforts is often required. Technology can help employees to be more effective and efficient in serving customers. Technology can also help customers become more educated and involved in cocreating service. In some cases, technology can serve as a substitute for employees, and actually deliver the service to the customer without any need for human interaction. These types of services – called *self-service technologies* – are prevalent today across industries.

A fourth difficulty associated with provider gap 3 involves the challenge in delivering service through such intermediaries as retailers, franchisees, agents, brokers, subcontractors, and outsourcers. Because quality in service often occurs in the human interaction between customers and service providers, control over the service encounter by the company is crucial, yet it rarely is fully possible. Most service (and many manufacturing) companies face an even more formidable task: attaining service excellence and consistency in the presence of intermediaries who represent them and interact with their customers yet are not under their direct control. Franchisers of services depend on their franchisees to execute service delivery as they have specified it. And it is in the execution by the franchisee that the customer evaluates the service quality of the company. With franchises and other types of intermediaries, someone other than the producer is responsible for the fulfillment of quality service. This issue has become particularly relevant as firms increase the outsourcing of their customer support to other countries; concerns about language, quality control, and consistency of performance are the ultimate trade-offs for lower costs. Because firms often provide service through intermediaries, they must develop ways to either control or motivate these intermediaries to meet company goals and perform as well as their own employees.

A final issue in provider gap 3 is the need in service firms to synchronize demand and capacity. Because services are perishable and cannot be inventoried, service companies frequently face situations of over demand or under demand. Lacking inventories to handle over demand, companies lose sales when capacity is inadequate to handle customer needs. On the other hand, capacity is frequently underutilized in service companies during slow periods. Most companies rely on operations strategies such as cross training or varying the size of the employee pool to synchronize supply and demand. Marketing strategies for managing demand – such as price changes, advertising, promotion, and alternative service offerings – can supplement approaches for managing supply.

Gap 4 – The Communication Gap

Even when a firm has done everything suggested by the other three gaps to ensure service quality, there can still be a failure to meet customer expectations if communications about the service do not match with what is delivered. Thus, the final provider gap that must be closed is the *communication gap*, or gap 4. This gap focuses on the difference between service delivery and what is communicated externally to customers through advertising, pricing, and other forms of communications.

Figure 6 captures several key strategies for closing gap 4. The first strategy revolves around

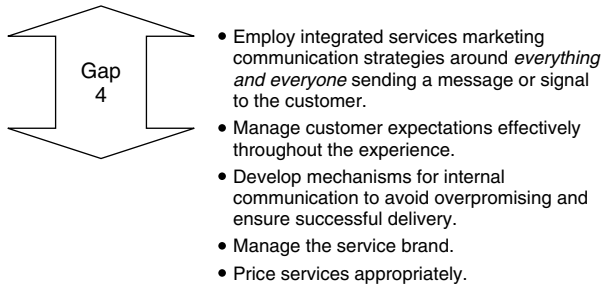


Figure 6 Strategies for closing the communication gap.

integrated services marketing communication that ensures that everything and everyone sending messages or signals about the service does so in a manner that is consistent with what customers expect and what is actually delivered. The challenge with this strategy is that there are a myriad of communication channels and modes that send messages to customers – more today than every before – including traditional websites, personal sales, direct mail, print media, blogs, virtual communities, mobile advertising, and television. Beyond these types of channels, which are also available to goods-producing firms, service customers receive additional communication from servicescapes, customer service representatives, and everyday service encounters with company employees. Ensuring that all of these channels communicate effectively and consistently is a daunting task, yet one that is essential to an integrated communication strategy. Unfortunately, the people within companies that deal with these different communication vehicles are not always located in the same department, leading to disparate, conflicting messages.

A second key strategy for closing the communication gap is to manage customer expectations effectively throughout the service experience. Many services (e.g., many business-to-business services and consumer membership services) take place over an extended time frame that might mean a few hours, days, weeks, or even years. These extended service experiences often change over time, varying from the original service promise as a result of business realities (for either the provider or the customer) that change the nature of the service, customer needs that change over time, and financial pressures that may cause increases in pricing or adjustments to the service contract. Thus, communications to the customer must also evolve through time to ensure that expectations and service performance match. This might mean managing customer expectations relative to new business realities, often in the form of managing expectations downward when a service previously provided is discontinued or when prices for similar services must be increased.

A third strategy for closing gap 4 is to develop mechanisms for internal communication so that the customer hears consistent messages before the sale and during service delivery. A common cause for the communications gap is overpromising on the part of customer-contact employees, salespeople, and marketing communications. While a certain amount of promotion is needed in many cases to gain a sale, excessive promotional activity can be detrimental when it exceeds the ability of the delivery organization to keep the promises made. Customers gained in the short term through excessive promises can be lost just as quickly through a failure to deliver. A number of internal communication strategies can help avoid the latter problem. These strategies including effective vertical communication that keeps employees informed of corporate plans and marketing messages so that they communicate accurately to consumers. Selling the brand inside the company also helps employees to see its value and to be realistic about what can and should be promised to customers. Horizontal communication across marketing, operations, and service design teams can also help align promises with service-delivery capabilities.

A fourth strategy to close the communications gap is to create a strong brand image for

the service. A brand image – reinforced by brand characters, slogans, logos, and similar images – serves to unify perceptions of the brand. Service organizations such as Chick-fil-a, with its ubiquitous cow, or McDonalds, with its golden arches, have successfully created brand images over time that reduce the communications gap.

A fifth issue in provider gap 4 is associated with the pricing of services. In packaged goods (and even in durable goods), customers possess enough price knowledge before purchase to be able to judge whether a price is fair or in line with competition. With services, customers often have no internal reference points for prices before purchase and consumption. For this reason, determining the value of a service to customers, called *demand-oriented pricing*, can be complicated. Further, pricing strategies such as discounting, "everyday low prices," and couponing obviously need to be different in service cases in which the customer has no initial sense of prices. Even cost-based techniques for developing prices for services are more complicated than those for pricing tangible goods, largely because the dominant cost factor is labor. Placing a value on an employee's time is more difficult than knowing the costs of components of a physical good.

Service Equals Profits

Through the 1980s and early 1990s, many firms jumped on the service bandwagon, investing in service initiatives and promoting service quality as ways to differentiate them and create competitive advantage. Many of these investments were based on faith and intuition by managers who believed in serving customers well and who felt in their hearts that quality service made good business sense. Indeed, a dedication to quality service has been the foundation for success for many firms, across industries. Since the mid-1990s firms have demanded hard evidence of the bottom-line effectiveness of service strategies. Fortunately, researchers are building a convincing case that service strategies, implemented appropriately, can be very profitable. Work sponsored by the Marketing Science Institute suggests that corporate strategies focused on customer satisfaction, revenue generation, and service quality may actually be more profitable than strategies focused on cost cutting or strategies that attempt to do both simultaneously (Rust, Moorman, and Dickson, 2002). Research from the Harvard Business School builds a case for the "service–profit chain," linking internal service and employee satisfaction to customer value and ultimately to profits (Heskett *et al.*, 1994). Furthermore, considerable research has shown linkages between customer satisfaction (often driven by service outcomes) and profits (Anderson and Mittal, 2000). The University of Michigan American Customer Satisfaction Index (ACSI) provides data suggesting that customer satisfaction is directly linked to shareholder value. Firms in the top 50% of the ACSI rankings show significantly higher shareholder value than do firms in the bottom 50% (Fornell, 2004). An important key to these successes is that the right service strategies are chosen and that these strategies are implemented appropriately and well.

Conclusion

Services marketing strategy focuses on delivering processes, experiences, and intangibles to customers rather than physical goods and discrete transactions. Delivering experiences successfully and building customer relationships are complicated undertakings involving many different strategies and tactics. Although companies have often found it difficult to attack service problems in an organized manner, a well-established model called the gaps model focuses on the customer and describes the approaches necessary to close the gap between customer expectations and perceptions. Figure 1, the full gaps model, shows that closing the all-important customer gap is a function of closing four gaps on the service provider side: the listening gap, the service design and standards gap, the performance gap, and the communication gap. Each of these gaps involves concepts and tools designed to minimize the gaps, and these were discussed in this article.

Bibliography

Anderson, E.W. and Mittal, V.S. (2000) Strengthening the satisfaction–profit chain. *Journal of Service Research*, 3, 107–120.

Bitner, M.O. (1995) Building service relationships: it's all about promises. *Journal of the Academy of Marketing Science*, **23**, 246–251.

Fornell, C. (2004) *Predictive Capabilities*, www.theacsi.org (accessed October 13).

Grönroos, C.N. (2007) *Service Management and Marketing: Customer Management in Service Competition*, John Wiley & Sons, Ltd, Chichester.

Heskett, L., Jones, T.O., Loveman, G.W. *et al*. (1994) Putting the service–profit chain to work. *Harvard Business Review*, **72**, 164–174.

Kotler, P. (1994) *Marketing Management: Analysis, Planning, Implementation, and Control*, 8th edn., Prentice Hall, Englewood Cliffs, NJ.

Parasuraman, A., Zeithaml, V.A., and Berry, L.L. (1985) A conceptual model of service quality and its implications for future research. *Journal of Marketing*, **49** (4), 41–50.

Rust, R.T., Moorman, C., and Dickson, P.R. (2002) Getting return on quality: revenue expansion, cost reduction, or both? *Journal of Marketing*, **66**, 7–24.

Vargo, S.L. and Lusch, R.F. (2004a) Evolving to a new dominant logic for marketing. *Journal of Marketing*, **68**, 1–17.

Vargo, S.L. and Lusch, R.F. (2004b), The four service marketing myths. *Journal of Service Research*, **6**, 324–335.

Zeithaml, V.A., Bitner, M.O., and Gremler, D.E. (2009) *Services Marketing: Integrating Customer Focus Across the Firm*, Mcgraw-Hill/Irwin, New York.

supply chain management strategy

John T. Mentzer and Terry L. Esper

Introduction

No firm is self-sufficient, as organizations have always relied on external trade partners for access to necessary resources. The contemporary realities of increasing competitive pressures, environmental uncertainties, and customer demands have heightened the strategic imperative for integration across functions and organizations in order to stabilize this resource access and exchange. As a result, many firms have begun to focus attention on supply chain management (SCM). *Supply chain management* is defined as "the systemic, strategic coordination of traditional business functions and the tactics across these business functions within a particular company and across businesses within the supply chain, for the purposes of improving the long-term performance of the individual companies and the supply chain as a whole" (Mentzer *et al.*, 2001, p. 18).

While the emerging definitions rightly consider SCM to be a cross-disciplinary concept, the fact remains that most academic research on SCM has focused on developing conceptual and theoretical frameworks that define what supply chains are, and exploring the processes associated with managing them. Noticeably lacking is extensive dialogue on the elements of SCM strategy. Academics and practitioners have primarily emphasized the execution of SCM processes, many times without appropriate strategic direction and alignment.

Generally speaking, the manifestation of a lack of strategic emphasis in SCM is cost inefficiency and/or customer service ineffectiveness. While the overarching paradigm in SCM has been to focus on efficiency and process optimization, this has often overshadowed the importance of managing supply chains effectively. In other words, the operations cost imperative has often taken precedence over the marketing customer service imperative. However, it is important to match supply chains with specific product characteristics and market conditions (i.e., Fisher, 1997; Lee, 2002). In essence, the underlying issue is ensuring that a firm's supply chain is constructed and managed in a way that effectively aligns with the needs and issues associated with customer demand and other environmental market forces, or, stated another way, ensuring that the supply chain is positioned for customer service effectiveness. This inherently involves purposively selecting and aligning supply chain exchange partners that are capable of supporting the strategic emphasis of the supply chain necessitated by product and market characteristics. As such, SCM strategy consists of aligning all companies involved in a supply chain with respect to product, market, and supply characteristics.

The key aspects of SCM strategy and some of the key facilitators of effective SCM strategy recognition and implementation are highlighted here. Toward that end, an overview of effectively defining the supply chain strategic emphasis is first discussed, followed by a discussion of supply chain design. The two fundamental facilitators of SCM strategy – demand and supply integration and supply chain orientation – are then described. Finally, concluding thoughts are offered, particularly emphasizing the marketing implications of SCM strategy.

SCM Strategy

SCM strategy involves two interrelated activities. To maintain acceptable levels of cost efficiencies and customer service effectiveness, firms must define the appropriate "goal" or strategic emphases of their supply chains. Although this is an important foundational principle of SCM strategy, it is not sufficient to arrive at a strategically relevant supply chain emphasis without the supply chain structure that supports the goal. Thus, in addition to strategic emphasis definition, SCM strategy involves effective supply chain design, whereby the structure of the supply chain is strategically architected.

The essence of SCM strategy lies in the fit between a firm's supply chain strategic focus and its structural support for effective SCM. The successful creation of a SCM strategy requires a high degree of *fit* between the identified emphasis and the supply chain structure. Fit is central to organizational success, and the best performing organizations establish strategies that closely fit the requirements of their environment

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

(Chandler, 1962). Further, the concept of fit has been extended to the supply chain environment, with a high degree of fit seen when member organizations align strategies in pursuit of common supply chain goals (Defee and Stank, 2005).

Defining the supply chain. Perhaps one of the most breakthrough ideas in evolving thought concerning supply chains is that they are not, and should not be, managed all the same. Given that products vary in demand patterns, and markets and competition vary in levels of stability, supply chains must be managed with a particular strategic emphasis. While all supply chains are, conceivably, different, there are different generic "types" of supply chains that suggest different strategic emphases.

Fisher (1997) distinguished between supply chains that should emphasize efficiency and those that should focus on responsiveness. According to Fisher, an efficiency-focused supply chain is one that places significant emphasis on cost reduction and inventory minimization. Responsive supply chains emphasize flexibility and responsiveness, often manifested through higher inventory safety stock levels and product customization. The criterion for defining the strategically appropriate emphasis is the nature of end products within the supply chain. Fisher contended that products can be generally characterized as either functional or innovative, where functional products are relatively low profit-margin products whose demand can be forecasted and variability predicted. Innovative products are characterized as having low demand predictability and often having higher profit margins, as they are typically new products or new variations of functional products. As such, Fisher suggested that the key to effective SCM is ensuring that efficiency-focused supply chains are aligned with functional products and responsive supply chains support innovative products. Hence, SCM effectiveness is based on the proper alignment of product characteristics and supply chain strategic emphases.

Lee (2002) provided a framework that essentially builds on the work of Fisher by adding the dimension of supply source stability. The Lee framework suggests four strategic emphases, two that are similar to the Fisher strategies, and two that deal specifically with supply environments that are relatively unstable. In particular, Lee suggests risk-hedging and agility as two strategic foci that emerge when the assumption of supply-side stability is relaxed, and relatively unstable supply markets are an issue. Again, the key idea is to define appropriate supply chain strategic perspectives depending on the nature of the external environment.

Goldsby, Griffis and Roath (2006) elaborated on this same concept in a manner quite similar to Fisher and Lee, although they referred to lean and agile supply chains. Furthermore, they discussed a postponement model as being "leagile," whereby the strategic emphasis is to leverage a hybrid approach, thus realizing benefits associated with both strategic emphases. This essentially entails operating with a lean emphasis until a certain point in the supply chain where operations become more agility focused. The key to "leagile" strategic operations is understanding when to most effectively convert from a lean focus to an agile focus. Again, this key juncture is often dictated by environmental characteristics, further supporting the notion of fit.

These strategic emphasis frameworks support the idea that supply chain strategy involves defining the appropriate strategic focus that should underscore SCM processes. Hence, supply chains emphasizing efficiency or "leanness" will have a certain manner in which inventory exchange and other supply chain exchange processes are managed and optimized, whereas supply chains more aligned with responsiveness should manage inventory investment, resource allocation, and exchange using different approaches. The ultimate goal is cost efficiency and customer service effectiveness; however, the strategically managed supply chain is one that emphasizes the appropriately defined means to that end.

Designing the supply chain. In addition to defining the strategic emphasis of a supply chain, it is important to design a supply chain that can effectively implement the defined strategic focus. While the full extent of the supply chain design concept is beyond the scope of the present discussion, the essence of supply

chain design can be captured by highlighting its core element – supply chain partner selection.

As previously mentioned, the supply chain concept has grown in popularity and ubiquity primarily because of the need for firms to maintain and manage access to necessary resources. As such, it is inevitable that firms will have to find resource exchange partners with which to engage. The key is to ensure that supply chain partners are capable of supporting and executing the defined strategic emphasis that environmental conditions necessitate. Two of the most important partner selection criteria are organizational compatibility and network structure, as they are conceptually broad enough to capture many of the issues that are associated with effective supply chain partner selection, and, ultimately, supply chain design.

Organizational compatibility. Organizational compatibility refers to the extent to which supply chain partners possess cultures, business objectives, managerial philosophies, and decision styles that are similar to, or complement, each other. The more similar or complementary these areas of the firms, the more likely a level of strategic agreement will exist that can support the necessary coordination of flows. This element is particularly important when firms are attempting to support the execution of a particular strategic emphasis across the supply chain. For example, when it is necessary to focus on responsiveness and agility, it will be difficult to appropriately exchange with a supply chain partner that has a very lean and efficiency-oriented culture and managerial philosophy.

Issues such as employee empowerment and organizational structure may also impede some of the strategic exchange efforts if the organizational representatives are not equally empowered or authorized to make strategic decisions related to the management of coordinated flows. Again, using the agility strategic emphasis as an example, it will prove challenging to execute such a strategy with supply chain partners that do not empower SCM personnel to make decisions in a timely manner. Agility requires a significant amount of change management and dynamism. As such, supply chain managers engaging in more agile or responsive exchange must be capable of making strategic-level decisions, often in the face of significant time pressure. Partner selection to support an agile strategic focus should, therefore, emphasize the degree to which representatives of potential partners are capable of engaging in agile and responsive decision-making.

Coordination of supply chain flows is not an easy task. The integration and synchronization of processes takes much time, effort, and investment. Furthermore, information sharing, which is such a fundamental feature of SCM, often requires a willingness to share sensitive information with supply chain partners. It is, therefore, important that there is sufficient managerial support within each partner firm to effectively invest the resources necessary to make coordinated flows a reality. This requires top managers to serve as champions of the coordination effort and provide organizational representatives with the level of empowerment, flexibility, span of control, and resource allocation necessary to support the intricacies of coordination in the supply chain.

In brief, organizational compatibility is about strategic alignment. In order for a defined supply chain strategic focus to be effectively implemented and executed, all supply chain partners must be strategically compatible. This essentially involves designing a supply chain that emphasizes alignment; ensuring that each partner has an internal infrastructure (i.e., culture, objectives, and managerial philosophy) that can effectively support a desired strategic emphasis within the supply chain. Effective partner selection should, therefore, involve assessments of such organizational artifacts when supply chain partners are under consideration.

Network structure. At the heart of supply chain partner selection is the creation of a relational agreement that will support significant resource exchange. While engaging in partner selection, supply chain managers must also consider the network implications of these relational agreements. In particular, it is important to be mindful of how selecting certain supply chain partners will impact the complexity and risk structure of the overall supply chain network. This is of particular concern in

contemporary supply chain structures, as global supply chains have become much more popular.

Geographic proximity, for example, is often an environmental element that can enhance integrated relationships. This issue is of significant importance when considering coordination in a global supply chain context. The ability to be located near each other is quite difficult when exchange partners are global. Technological innovations such as teleconferencing and simple e-mail communication have significantly reduced this barrier. However, close geographic proximity is still an important issue in coordinating flows with supply chain exchange partners. Many large organizations often require key customers and suppliers, with whom coordination is extremely necessary, to develop satellite locations in close proximity to their facilities. This allows for more seamless coordination and exchange flows, as well as easy access to customer/supplier support personnel.

Overall, the issue is managing complexities and risk in the supply chain network structure. This is of particular importance when one considers the different strategic emphases that govern supply chains. For example, a prevalent difficulty is managing lean processes in global supply chains. The reduction in geographic proximity associated with managing global supply chain exchange has often increased lead times, managerial complexities, and operational risks. Such complexities and risks are very difficult to justify when attempting to focus on lean and efficient operations. Hence, many supply chain managers have opted to reduce global partner selection efforts to effectively support the lean strategic supply chain emphasis. The key is selecting partners that allow for a supply chain network that effectively supports the appropriate strategic emphasis.

When considered in concert, SCM strategy involves assessing the characteristics of the supply and demand environments to define a strategic emphasis that allows for optimal customer service effectiveness, and design a supply chain structure that can appropriately execute. This involves the notion of "fit" aligning environmental forces with the focus and network of exchange management along the supply chain. Hence, supply chain managers should determine if the supply chain should emphasize efficiency, responsiveness, risk-hedging, or a hybrid approach such as a "leagile" focus, then design the appropriate supply chain by effective and aligned partner selection and network structuring.

Facilitators of SCM Strategy

In order to successfully engage in SCM strategy, there are certain necessary facilitators. A significant impediment to thinking about the supply chain strategically is a lack of emphasis on SCM within firms, and a lack of effective communication across the organization. Firms that are more often "best in class" in managing supply chain processes have an internal infrastructure that supports strategic SCM thinking and execution. This essentially involves open lines of communication and exchange between the demand and supply sides of the firm, and an overall internal orientation that places strategic value on the role of SCM in supporting the firm's competitive positioning in the marketplace.

Demand and supply integration. As depicted in Porter's (1985) value chain framework, there are two primary sets of processes through which a firm creates value for its customers by moving resources and information along the supply chain: demand-focused processes consisting of marketing, sales, and customer relationship management activities; and supply-focused processes consisting of inbound logistics, operations, and outbound logistics. Historically, firms have invested resources to develop a core differential advantage in one of these areas – but rarely in both – often elevating the importance of one area at the expense of the other. The manifestation of such an approach is what Drucker (1973) has referred to as "The Great Divide" – separating the processes used to plan for and manage customer demand from those required for supplying the resources and operational dexterity to meet that demand.

Demand-focused firms tend to create value through an emphasis on effectiveness in serving customer needs at the expense of efficiency, whereas supply-focused firms tend to create value through an emphasis on efficiency at the expense of effectiveness. Too often, however, the traditional isolation of demand

and supply processes results in customer service ineffectiveness, manifested through mismatches between demand (i.e., shortages of products that customers want and/or surpluses of products that are not wanted), and supply (i.e., what is actually available in the marketplace).

Recently the concept of "demand and supply integration" (DSI), the balancing of demand and supply market information and business intelligence through integrated knowledge management processes to strategically manage demand and supply activities for the creation of superior customer value (Esper *et al.*, 2010), has received attention. This essentially entails a joint consideration of demand and supply-side information and processes to make more informed supply chain decisions throughout the firm. By simultaneously considering applicable market developments from both upstream and downstream forces, organizations are poised to exploit and leverage efficiency-focused operations while maintaining relevant levels of effectiveness.

An application of the DSI concept is sales and operations planning (S&OP), a tactical-level process for balancing and matching supply and demand plans. The S&OP process often focuses on making decisions about how to enhance demand when supply exceeds demand (e.g., increased advertising expenditures, pricing adjustments, new promotional activity), or how to dampen demand when demand exceeds supply capacity (e.g., reducing advertising, raising prices, discontinuing promotional activity, and incentivising customers to switch to other products and services). Again, the focus is on taking a balanced and comprehensive approach to managing supply and demand forces and processes.

DSI is, therefore, an important facilitator of SCM strategy because it supports the type of communication and strategic consideration that is necessary to make informed strategic SCM decisions. To reiterate, the core of SCM strategy is designing and managing operations within the supply chain that fit both demand and supply-side market characteristics. Firms that suffer from lack of integration run the risk of designing and managing their supply chain in a way that does not fit their environmental structure. For example, very supply-focused firms often emphasize "leanness" and efficiency in managing their supply chain. Yet, without involvement from the demand side of the firm, it is possible that a more responsive supply chain is required based on market demand unpredictability. Likewise, demand-side orientation may result in an emphasis on responsiveness and agility, where in actuality the level of supply-side stability dictates more of an emphasis on lean operations. Thus, it is difficult to ensure an aligned strategic supply chain definition and design if there is not a level of integration to support the knowledge of what type of supply chain both the supply and demand environments necessitate.

In addition to DSI, the firm must embrace the SCM concept. Even if a firm effectively engages in DSI, and allows for the joint consideration of demand and supply market information to make decisions, it does not automatically mean that the firm will leverage this information for SCM. An additional facilitator of SCM strategy is an organizational infrastructure and philosophy that supports SCM.

Supply chain orientation. Supply chain orientation (SCO) is defined as "*the recognition by a company of the systemic, strategic implications of the activities and processes involved in managing the various flows in a supply chain*" (Mentzer *et al.*, 2001, p. 14). Hence, while SCM focuses on the management of exchange flows within and across the members of the supply chain, SCO emphasizes the strategic awareness and embracing of SCM within an individual supply chain firm. SCO is a necessary antecedent to effective SCM, which suggests that an organization must first look inward before it can effectively engage in strategic management of supply chain processes (Min and Mentzer, 2004).

The underlying premise of the SCO concept is the notion that effective implementation of SCM requires the firm to place strategic emphasis on the internal infrastructure, or culture, necessary to facilitate SCM exchange. In essence, SCO represents a shared value and belief system that aids in understanding how the organization should strategically manage its supply chain, and the behavioral norms needed inside the organization (Deshpande and Webster, 1989). This suggests that organizations possessing an SCO

approach SCM differently from firms that are less inclined to view SCM strategically.

SCO has been viewed from both strategic and structural perspectives. The strategic SCO paradigm builds on the original conceptualization of the construct, where SCO was conceived as a philosophy focused on the implications of managing supply chain flows (Mentzer *et al.*, 2001). It conceptualizes SCO through an emphasis on the importance of strategic direction in managing supply chains. As such, the nature of this perspective involves making a strategic choice to compete on the basis of supply chain capabilities (Defee and Stank, 2005) and utilizing this strategic emphasis to drive the performance of strategic business units within the firm (Stank *et al.*, 2005). The strategic SCO perspective encourages firm personnel to act in a manner that manages flows from supplier to customer, taking a systems approach to viewing the supply chain holistically rather than as constituent parts, and seeking integration, synchronization, and convergence of intrafirm and interfirm operational and strategic capabilities (Min and Mentzer, 2004).

The structural SCO perspective, on the other hand, emphasizes organizational artifacts that facilitate SCM. For example, Min, Mentzer and Ladd (2007) suggested that SCO involves building and maintaining internal behavioral elements that facilitate relational exchange. Mello and Stank (2005) focused on defining SCO as cultural elements that support the organization's structure. They suggested that supply chain oriented organizations should behave in a way that facilitates trust, commitment, compatible supply chain partners to cooperate with, and top management endorsement. Though not referring to as SCO per se, Trent (2004) highlighted intrafirm structural elements necessary for effective SCM. This perspective focuses on an intrafirm structural management approach that facilitates effective SCM through emphasis on the behaviors, systems, and cultures necessary for integrated supply chain exchange.

Considering both perspectives leads to the conceptualization that a supply chain– oriented firm not only places strategic emphasis on systemic, integrated SCM, but also aligns this strategic thrust with an intraorganizational structure that capitalizes on this strategy. The strategic perspective alone is not enough, because it requires structural support for effective implementation. Furthermore, the structural perspective does not, in and of itself, provide the necessary strategic direction to undergird the development and management of the structure. Hence, both perspectives seem to capture the true essence of the SCO concept.

Conclusion

The key elements of SCM strategy are to define a strategic emphasis for the supply chain, and to design the appropriate upstream and downstream supply chain to support this emphasis. This entails developing a vision of the supply chain to meet the market and supply characteristics faced, selecting appropriate supply chain partners, and building an efficient and effective network structure.

This is difficult to implement without the definitional and design influences of DSI and an SCO. DSI provides the orientation and processes to match market demand with supply chain capacities. SCO provides a cultural awareness that thinking about the supply chain strategically is of relevance and importance. One of the primary impediments to strategic SCM is that the firm does not value SCM, that is, its strategic role has not been defined. Hence, a key facilitator of SCM strategy is the development of an internal awareness and appreciation for SCM and how leveraging SCM processes for competitive positioning contributes to effective marketplace differentiation.

Many executives cannot cite a consistent, or even well-thought-out, definition of what SCM means in their company. It is perhaps axiomatic to say that design cannot be accomplished if what is being designed has not been first defined. Thus, companies trying to pursue a competitive advantage through SCM must first define their market and supply characteristics and related DSI implications, define what supply chain these characteristics necessitate, design the supply chain that will meet these definitions, and then select the supply chain partners to pursue this strategic vision.

Of course, the tactics of relational contracting, trust building, performance assessment, and reward sharing, and regular communication

between supply chain partners is critical, but these are more in the realm of the management of the supply chain, given an overarching SCM strategy. This, of course, leads to another vital managerial implication – companies successful at supply chain strategy and management must constantly reassess the factors mentioned in the previous paragraph and make certain they are still in concert with the managerial implementation of the latter factors.

The definitions open a fascinating vista of questions to be investigated concerning the definition, design, and management of supply chains. However, more work is needed to define the domains and dimensions of each of the theoretical constructs that precipitate from the discussion above. What broader, extant theories can be brought to bear on this definitional effort? This also leads to work on developing measures of each perspective. Subsequently, a myriad of questions present themselves. Does this process of definition, design, and implementation affect supply chain performance overall, and the performance of the specific firm with the SCM strategy and its partners? Are there significant differences in the distribution of rewards between these two groups, and are these differences seen as "unfair" by supply chain partners? Does the nature of the definition, design, and implementation activities vary across global cultures, and if so, does it affect supply chain strategy and management? Answers to these questions, and many other questions that will result from future thinking, should provide a rich body of knowledge for both studying SCM and using SCM for competitive advantage.

Bibliography

Chandler, A.D.J. (1962) *Strategy and Structure*, The M.I.T. Press, Cambridge.

Defee, C.C. and Stank, T.P. (2005) Applying the strategy – structure – performance paradigm to the supply chain environment. *International Journal of Logistics Management*, **16** (1), 28–50.

Deshpande, R. and Webster, F.E. Jr. (1989) Organizational culture and marketing: defining the research agenda. *Journal of Marketing*, **53** (1), 3–15.

Drucker, P.F. (1973) *Management*, Harper & Row, New York.

Esper, T.L., Ellinger, A.E., Stank, T.P. *et al.* (2010) Demand and supply integration: a conceptual framework of value creation through knowledge management. *Journal of the Academy of Management Science*, **38** (1), 5–18.

Fisher, M.L. (1997) What is the right supply chain for your product? *Harvard Business Review*, **75** (2), 105–116.

Goldsby, T.J., Griffis, S.E. and Roath, A.S. (2006) Modeling lean, agile, and leagile supply chain strategies. *Journal of Business Logistics*, **27** (1), 57–80.

Lee, H. (2002) Aligning supply chain strategies with product uncertainties. *California Management Review*, **44** (3), 105–119.

Mello, J.E. and Stank, T.P. (2005) Linking firm culture and orientation to supply chain success. *International Journal of Physical Distribution and Logistics Management*, **35** (8), 542–554.

Mentzer, J.T., DeWitt, W., Keebler, J.S. *et al.* (2001) Defining supply chain management. *Journal of Business Logistics*, **22** (2), 1–25.

Min, S. and Mentzer, J.T. (2004) Developing and measuring supply chain management concepts. *Journal of Business Logistics*, **25** (1), 63–99.

Min, S., Mentzer, J.T. and Ladd, R.T. (2007) A market orientation in supply chain management. *Journal of the Academy of Marketing Science*, **35** (4), 507–522.

Porter, M.E. (1985) *The Competitive Advantage: Creating and Sustaining Superior Performance*, Free Press, New York.

Stank, T.P., Davis, B.R. and Fugate, B.S. (2005) A strategic framework for supply chain oriented logistics. *Journal of Business Logistics*, **26** (2), 27–45.

Trent, R.J. (2004) What everyone needs to know about SCM. *Supply Chain Management Review*, **8** (2), 52–59.

SWOT analysis

Srini Srinivasan

Business scholars and managers have long been interested in understanding the sources of competitive advantages of firms (Barney, 1995). Researchers at Harvard Business School, investigating the drivers of organizational performance, formalized SWOT (strengths, weaknesses, opportunities, and threats) analysis as a strategic tool in managing organizations for better performance (Learned *et al.*, 1965). SWOT analysis posits that to compete successfully in the market place, organizations need to consider concurrently both internal and external factors – some of which may be favorable to the organization while others may not (Valentin, 2001). In the SWOT analysis, internal factors that drive firm performance are broadly classified into strengths and weaknesses. Businesses derive their strengths from superior skills and tangible and intangible resources that help them to perform better in the market place and thwart competition. Weaknesses are deficiencies and constraints that prevent a firm from performing optimally or succeeding in the market place (Day, 1990). SWOT analysis also places emphasis on identifying relevant external factors that impact firm performance, and it broadly classifies them into opportunities and threats. Opportunities are trends and events outside the business that can significantly improve firm performance if appropriate actions are taken by the organization. Threats are environmental factors, which, in the absence of proper strategic response, are likely to result in significant decrease in firm performance and erode its competitive advantage (Aaker, 2005). A thorough SWOT analysis helps firms to build strategies that utilize their strengths and neutralize their weaknesses.

Though SWOT analysis was conceptualized more than four decades ago, it is still popular among executives owing to its simplicity and ease of use (Piercy and Giles, 1989). It is considered useful as it provides a simple framework for managers to brainstorm and map how they perceive their organization and their business environment. However, some researchers feel that the open nature and unstructured method of SWOT analysis offer little guidance to users on where to search for the variables or what to do after identifying them (Panagiotou, 2003) and it is likely to identify too many factors, which have only minimal impact on firm performance. SWOT analysis is criticized for its static and descriptive nature and its inability to provide any analysis of the impact of these factors on firm performance. Weaknesses identified in the SWOT analysis can be overcome with additional resources, if suitable opportunities, which make the investment worth it, are present in the environment. The opportunities and threats identified are not static and sometimes threats appear in the guise of opportunities or vice versa. To improve the usefulness of SWOT analysis, Hill and Westbrook (1997) suggest that researchers provide more guidance on selecting the factors, and prioritizing them. Since the 1960s, when SWOT analysis was originally proposed, business disciplines have made strides in understanding firm performance and the factors that drive success (Day, 1990). SWOT analysis is likely to continue to be valuable in future, and business executives are likely to use it in conjunction with other strategic tools to create and extract value in the market place.

Bibliography

Aaker, D.A. (2005) *Strategic Market Management*, 7th edn, John Wiley & Sons, Inc.

Barney, J.B. (1995) Looking inside for competitive advantage, in *Academy of Management Executive (1993-2005)*, Vol. 9, Academy of Management, pp. 49–61. http://www.jstor.org/stable/4165288.

Day, G.S. (1990) *Market Driven Strategy: Processes for Creating Value*, Free Press, New York.

Hill, T. and Westbrook, R. (1997) Swot analysis: it's time for a product recall. *Long Range Planning*, **30** (1), 46–52.

Learned, E., Christensen, C.R., Andrews, K.R. and Guth, W.D. (1965) *Business Policy: Text and Cases*, Richard D Irwin, Homewood.

Panagiotou, G. (2003) Bringing Swot into focus. *Business Strategy Review*, **14**, 8–10.

Piercy, N. and Giles, W. (1989) Making Swot analysis work. *Marketing Intelligence and Planning*, **7** (5), 5–7.

Valentin, E.K. (2001) Swot analysis from a resource-based view. *Journal of Marketing Theory and Practice*, **9** (2), 54–69.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

trademarks, proprietary marks, and brands

Philip C. Zerrillo

A *trademark* includes any word, name, symbol, device, or any combination, used, or intended to be used, in commerce to identify and distinguish the goods of one manufacturer or seller from goods manufactured or sold by others, and to indicate the source of the goods. In short, a trademark is a brand name.

The trademark is often equated with its historical precursor, the regulatory production mark. The regulatory production mark has its origins in the crafts guilds of Europe, most commonly believed to have started with the cooper's guild. Craftsmen would affix their mark to the inside of the trunks they produced as a means of identifying the work of their craftsman or guild.

A trademark can be registered, and become a "registered trademark" when it is filed with a government licensing office. (i.e., in the United States it is the US Patent Office) and granted registered status. These registered marks will then carry a ® to denote their legally registered status.

The *proprietary marks*, unlike the regulatory production mark, do not designate who the makers of the product are but rather who the owners of the goods are. In cases of shipwreck or piracy, the proprietary mark served as a useful means for sorting and recovering the dislocated goods. In the southwestern United States the cattle industry has a long and storied history with proprietary marks as the cattle were "branded" by burning the hide with a proprietary mark. This mark did not signify the maker or even the animal's breeder, but rather the owner. These marks did allow for the cattle to be sorted and recovered when pasture fences were broken and the herds were mixed, or when a rustler stole the herd.

A *brand*, as defined by the American Marketing Association (AMA), is a distinguishing name and/or symbol (such as a logo, trademark, or package design) intended to identify the goods or services of either one seller or a group of sellers, and differentiate those goods or services from those of competitors.

From a legal standpoint, a brand thus signals to the customer the source of the product and protects both the customer and the producer from competitors who would attempt to provide products that appear to be identical. The European and North American courts initially developed laws to protect the brand names and marks from would-be infringers as a means to protect consumers from the tort of fraud. In other words, the primary concern was to keep the consumer from being taken advantage of by producers that would tread on the good name of well-known manufacturers to pass off inferior goods. Over the past 70 to 90 years the emphasis of the law has swung toward not just protecting the consumer from fraud, but the protection of the business assets that brands represent (Peterson, Smith, and Zerrillo, 1999). In practice, from a legal standpoint, the emphasis of brand protection has begun to model property law.

The managerial emphasis of branding has been to move beyond commodities to branded products in order to reduce the importance of price in the purchase decision by accentuating the basis of differentiation. In other words, a brand tells us whose product it is, and whose product it is not, so that we may differentiate the goods from those of other producers and provide a basis for loyalty. But at the end of the day, the goal is to get away from price as the reason to buy a good or service. For instance, in the absence of any other information about two bags of cement, say cement A and cement B, the purchaser will normally buy the cement that costs less, as it is a way of reducing their risk. But, when you put cement A in a shiny bag, and inform would be buyers that it is fast drying, has a 30-year life expectancy, and can be submerged in 200 meters of water, the consumer may be willing to pay more for A than B. Thus, the ability to differentiate the offering and promote features and performance rather than price is a means of moving beyond commodities.

The value of brands are widely studied and discussed. They often dwarf the value of plant and equipment when firms and their brands are purchased. (*see* PERCEPTION OF BRAND EQUITY). Investment in brand-building activities is often a struggle within the modern organization. On the one hand, the brand is a long-term asset and it is easy to temporarily cut expenditures on brand building activities such as sales, advertising, and trade promotion voluntarily as a

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

means to increase contribution margins. Moreover, the struggle within firms often centers around the realization that while brands have great value (*see* BRAND VALUE), the activities that build them (advertising, sales-force promotion, channel management) are treated as operating expenses rather than as investments by the finance department.

Bibliography

Peterson, R., Smith, K. and Zerrillo, P. (1999) Trademark dilution and the practice of marketing. *Journal of the Academy of Marketing Science*, **27** (2), 255–268.

a framework for creating value propositions

James C. Anderson and Gregory S. Carpenter

Definition

What is a customer value proposition, or simply, a value proposition? A *value proposition* is the statement of those few points of difference that create a compelling reason for a customer to select one market offering relative to competing alternatives.

Framework for Creating Value Propositions

The three primary factors in constructing value propositions that target customers find persuasive are frame of reference, points of parity, and points of difference. The *frame of reference* is the set of potential competing alternatives that a customer considers. *Points of parity* are the elements of the market offering on which the customer views the offerings as essentially the same. *Points of difference* are the elements of the offering on which the customer views one offering as noticeably superior.

Potential value propositions can be evaluated on three criteria: distinctiveness, measurability, and sustainability (Anderson, Kumar, and Narus, 2007). *Distinctiveness* captures the extent to which the value proposition is uniquely able to deliver on a claim relative to alternative offerings. *Measurability* is the extent to which the value proposition can be substantiated in tangible ways, such as in monetary terms. *Sustainability* captures how long the uniqueness the value proposition conveys is expected to last.

Value Propositions for Consumer Markets

In consumer markets, one way to understand value is to consider that customers seek to achieve goals. Goals take two forms: positive outcomes we hope to achieve and negative outcomes we hope to avoid. Achieving their goals creates a problem for buyers to solve. Market offerings create value for buyers by proposing solutions to their problems. Royal Caribbean cruises offer one example.

To solve problems, buyers often consider a number of alternatives. On vacation, some might seek relaxation, others seek a cultural experience, and still others might seek to indulge a personal goal (learn Italian). Royal Caribbean identified a group of consumers who seek adventure on vacation. They call this group "explorers." For explorers, vacation options include helicopter skiing, scuba diving, or bicycle touring.

Although, all these options share the element of adventure, each option offers a distinctly different experience. In comparison, a cruise is seen as sedate, unexciting, and for older people. A cruise, however, does have one significant point of difference with other vacations – it offers variety in terms of locales and activities.

To craft a value proposition, Royal Caribbean developed the idea that variety can be a dimension of adventure – a cruise offers *multiple* new experiences that are not possible with a single-adventure holiday. By offering multiple novel experiences, a Royal Caribbean cruise can offer a greater sense of adventure.

The value proposition they developed reflects this essential insight: Royal Caribbean is the vacation option that creates unexpected experiences that brings out the explorer in me.

As a result of this new value proposition, brand awareness and preference have increased among target buyers, the customer mix is increasingly younger than other cruise lines, and booking has been growing.

Value Propositions for Business Markets

Management practice research (Anderson, Narus, and van Rossum, 2006) has recommended that suppliers in business markets construct *resonating focus* value propositions, which consist of "the one or two points of difference (and, perhaps, a point of parity) whose improvement will deliver the greatest value to the customer for the foreseeable future" (p. 93). The rationale for resonating focus value propositions is that in a business world where customer managers are taking on greater responsibility and are increasingly pressed for time, to be successful, suppliers must deliver customer value propositions that are simple, yet powerfully captivating. They do this by making their

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

offerings superior on the few elements whose functionality or performance matter most to target customers, demonstrate and document the value of this superior performance, and communicate it to target customer managers in a way that conveys to them that the supplier understands their business concerns and priorities.

A resonating focus value proposition is preferable to more commonly used value propositions, which consist of either all benefits that the customer receives from a market offering or all favorable points of difference that an offering has relative to the next best alternative.

Bibliography

Anderson, J., Kumar, N., and Narus, J. (2007) *Value Merchants: Demonstrating and Documenting Superior Value in Business Markets*, Harvard Business School Press, Boston.

Anderson, J., Narus, J., and van Rossum, W. (2006) Customer value propositions in business markets. *Harvard Business Review*, **84**, 90–99.

marketing warfare strategies

Daniel Laufer

Marketing warfare strategies are competitor-centered strategies borrowed from the field of military science and applied in a marketing context. According to Rindfleisch (1996), advocates of this approach suggest that both marketing and the military are very similar, and both fields share "(i) two or more opposing forces; (ii) competition for a limited and valuable resource; (iii) a zero-sum orientation in which gains come at the expense of competitor losses; and (iv) defined rules of engagement in which one competitor attacks while the opposing defends." This approach was popular in the 1980s, whereas in the current era of relationship marketing this approach has become less relevant. Also, an increased awareness of differences between business and military organizational cultures raises questions of the applicability of military strategies in a business context.

In 1981, Kotler and Singh (1981) wrote a highly influential article about how warfare strategies can be applied in the field of marketing. This competitor-centered approach suggested that principles from military strategy can be applied to develop (i) business objectives; (ii) attack strategies; and (iii) defense strategies.

Kotler and Singh (1981) claimed that the concept of "principle of objective" derived from military science can also be applied to business objectives. The principle of the objective stated that "every military operation must be directed toward a clearly defined, decisive, and attainable objective" (Kotler and Singh, 1981 p. 33). According to Kotler and Singh (1981), in a marketing context the objective is usually described in terms of "crushing the competitor, reducing its share, or freezing its share" (Kotler and Singh, 1981 p. 33), and is contingent on the type of competitor the company is facing (market leader, similar-sized firm, or smaller firm).

After determining the objective, a company assesses how to wage war. Kotler and Singh (1981) introduced two types of strategies: attack strategies and defense strategies. Attack strategies involve taking the offensive and seizing market share from competitors. These strategies include frontal attacks, flanking attacks, encirclement attacks, bypass attacks, and guerrilla attacks (Kotler and Singh, 1981). Defense strategies, on the other hand, involve fortifying one's position, or defending market share from competitors. These strategies include position defense, mobile defense, preemptive defense, flank-positioning defense, counteroffensive defense, and strategic withdrawal (Kotler and Singh, 1981).

Interest in marketing warfare strategies peaked in the 1980s during an era of slow economic growth. The economic downturn caused companies to seek growth opportunities primarily from gains in market share as opposed to increases in overall market demand. This encouraged companies to turn to competitor-centered strategies such as marketing warfare strategies. In addition, prominent military officers moved into key leadership positions during this time at business-related organizations which influenced the choice of corporate strategy due to the military experience of the new executives (Rindfleisch, 1996). Finally, large corporations turned to the military for guidance in creating an effective bureaucracy to run their organizations. This created similar organizational structures and policies and procedures, such as chains of command and standard operating procedures (Rindfleisch, 1996). Similarities between the bureaucracies encouraged a military mindset among business organizations which generated increased interest in marketing warfare strategies

During the 1990s and 2000s, marketing warfare strategies became less relevant during an era of relationship marketing. A number of researchers questioned whether the business context is equivalent to a war despite the similarities on a number of dimensions (Rindfleisch, 1996). In particular, a zero-sum game mentality regarding winners and losers became less prevalent, and cooperation and coexistence among competitors became more the norm in many industries. In addition, there was an increased awareness during this time period of differences between the organizational cultures of businesses and the military. As Rindfleisch

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

(1996) points out, not all businesses are highly structured and "the type of highly structured and calculating strategic planning implied by marketing-as-warfare may be more fitting for companies with bureaucratic cultures and less appropriate for firms with more clan-like cultures" (Rindfleisch, 1996 p. 7).

Bibliography

Kotler, P. and Singh, R. (1981) Marketing warfare in the 1980's. *Journal of Business Strategy*, **1** (Winter), 30–41.

Rindfleisch, A. (1996) Marketing as warfare, reassessing a dominant metaphor. *Business Horizons*, **39**, 3–11.

causal research

Harmen Oppewal

Causality concerns cause-and-effect relationships between variables (*see* CONCEPT OF CAUSALITY AND CONDITIONS FOR CAUSALITY). Humans have a natural tendency to think in terms of cause-and-effect relationships and we use such relationships to "explain" the world we live in. As individuals, we continuously observe and monitor changes in our environment and infer from these changes what the likely cause is for some outcome. Knowing about causes is crucial for taking decisions about what actions to take, from small day-to-day decisions ("why am I feeling cold? – is the heating system on?") to the most important issues in society ("is there global warming and what are its causes; will reduction of carbon dioxide emissions help reduce global warming?"). Whereas philosophers have debated for centuries whether causality exists, people use it, and rely on it.

The scientific process of generating, testing and revising ideas can be seen as an extension and refinement of the basic human processes of knowledge generation. In this, testing which relationships are causal and which are "spurious" is seen as one of the main aims of scientific research, both by scholars who aim to find the "truth" about what influences and determines observed phenomena, and by scholars and practitioners who are keen to just know "what works" (Dunbar, 1995). As an example in marketing, consider the role of satisfaction (*see* CUSTOMER-SATISFACTION RESEARCH). It is intellectually gratifying to know what determines satisfaction, but it is also very relevant practically for marketers to know what actions can be taken to increase satisfaction. Should a company invest in a loyalty card scheme in order to increase customer satisfaction and retention? Such investments will only be worthwhile if the scheme can be expected to indeed positively affect satisfaction and retention rates. Causal research is relevant even if it would lead to the conclusion that such schemes are *not* contributing, as in that case the company can better use its resources for other purposes. Identifying the cause of some outcome can be as important as identifying if some factor is not a cause.

Causal research is research that aims to investigate causal relationships. Causal research therefore always involves one or more independent variables (or hypothesized causes) and their relationships with one or multiple dependent (or outcome) variables. In addition there may be "mediating" variables sitting between the independent and dependent variables, reflecting the causal process or mechanism. Causal *descriptive* research aims to establish the *effect* of the independent variable(s) on the dependent variable. Experiments are widely accepted as the best method for testing such effects, and the remainder of this article will focus largely on experiments. However, knowing an effect exists and its size is not equal to having insight as to why the effect exists. Causal *explanatory* research aims to describe the process of how an independent variable influences an outcome variable (Shadish, Cook, and Campbell, 2002). Explanatory research requires theorizing and the use of abstractions, or constructs, to identify why outcomes occur the way they do. Causal explanatory research can include the generation of new ideas and insights on the basis of literature research, participatory observation (*see* ETHNOGRAPIC RESEARCH), focus groups (*see* FOCUS GROUPS AND DEPTH INTERVIEWS), and any other type of data collection, and can also use analogies or models from other domains. In science, there is no limit and restriction to idea generation, but to become an accepted theory the ideas have to fulfill basic requirements such as coherence and consistency, and the ability to explain phenomena that we have accepted (and that we are willing to continue to accept) as "facts" (Dunbar, 1995; Latour, 1987). This wider area of theory building and model specification will not be further discussed in this article but note the issue of *validity*, especially construct validity (*see* RESEARCH RELIABILITY AND VALIDITY; CONCEPT OF CAUSALITY AND CONDITIONS FOR CAUSALITY).

A particular and more formalized type of causal explanatory research concerns "causal modeling." Causal modeling tends to be associated with *structural equation modeling* (SEM), which extends regression and factor analysis to

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

allow the measurement of "latent" constructs and their causal relationships within one statistical framework (*see* STRUCTURAL EQUATION MODELING). In SEM the assumed causal relationships between constructs are displayed in a "path diagram," which is a set of boxes and arrows showing how the constructs influence each other. Causal modeling can also be considered more widely as to include the various regression-based models used in econometrics that capture the relationships between independent and dependent variables in various types of statistical formulations (*see* MULTIPLE REGRESSION). The key feature of all these methods, whether SEM or econometric models, is that they are typically applied to data that are either cross-sectional or, if time series are used, lack a proper control condition. They capitalize on the use of statistical theory and mathematical functions to determine and disentangle the effects of independent variables, and the theoretical constructs they represent, on dependent variables (or constructs). Whereas they are very sophisticated and potentially very useful for assessing such relationships, their application to cross-sectional, uncontrolled data shows that they cannot fully establish or test the causal nature of relationships in the data. They can only attempt to *statistically* separate possible causes from noncauses. In contrast, experiments aim to separate and control various causes *by design*. This is why conclusions from experiments typically have greater *internal validity*; they are a better basis for establishing whether the independent variable was the true cause of the effect observed for the dependent variable. Experimental control, however, cannot always be achieved and further, has often the disadvantage of limited external validity. In this respect, the design and statistical approaches complement each other and in practice are often used together to maximize the validity of the knowledge gained from a research project (*see* VALIDITY IN EXPERIMENTATION).

Principles of Experimental Design

Basic terminology. Although, many people associate scientific experiments with laboratories, white coats, and guinea pigs, experiments have a much wider use. Experiments are sophisticated versions of trial and error, or plan-act-evaluate cycles of behavior: we try something, observe the outcome, and draw conclusions regarding the effects of our actions.

We now turn to discussing the key concepts and rationale of experimentation and experimental design. Before discussing various types of design it is useful to introduce the notation system to graphically display the characteristics of the various designs and discuss the principles of design and analysis (Cook and Campbell, 1979).

Each test unit in the experiment undergoes a sequence of events consisting of *treatments* (X), representing variations of the independent variable, and produces *observations* (O), which are measurements of the dependent variable. Figure 1 displays three *preexperimental* designs. They each include a treatment group and one or more observations. Neither however has a comparison or control group to act as a benchmark for assessing the effects of the treatment. These designs therefore cannot automatically rule out the "third variable" alternative explanations that threaten the internal validity of conclusions drawn from these designs (*see* VALIDITY IN EXPERIMENTATION). Suppose we conduct a study on advertising effectiveness. X represents "being exposed to an advertisement" and O is a participant's response to a brand perception question (Figure 1a). If O is positive, can we infer that the advertisement exposure caused the positive effect? Not so, as many other variables will (also) have influenced the participant's response. We could include a pretest measure as in Figure 1b. This "one-group pretest–posttest design" would allow observing how the brand perceptions have changed since the advertisement was presented, for example, if $O1$ and $O2$

Preexperimental designs

a. One shot case study

X O1

b. One-group pretest-posttest design

O1 X O2

c. One-group time-series experiment

O1 O2 O3 X O4 O5 O6

Figure 1 Preexperimental designs.

had been observed with one week interval. However, this would not separate the advertisement effect, if it existed, from the possible effects of other events occurring between $O1$ and $O2$. Even if a longer time series (Figure 1c) had been observed, although providing a stronger basis for inferring the effect of X on O, it would still not account for events co-occurring with X.

Experimental designs include at least two groups (conditions), one being the experimental group and the other the control group. The *control group* does not receive the experimental treatment. An example is shown in Figure 2. G1 could be a group seeing the advertisement and then being asked to rate our brand, while G2 could be a group that does not see the advertisement, but still rates the brand. Comparing $O1$ and $O2$ will give an indication of the effect of X. However, there may be preexisting differences between G1 and G2, and these may be the underlying causes of observing differences on the values of $O1$ and $O2$. To overcome this issue experimental designs rely on *randomization* (R). Experiments that use existing groups without randomization are called *quasi-experiments*.

Figure 3 presents a two-group *posttest only* experimental design. It is a posttest only design because the measures are only taken after the treatment (seeing the advertisement) has been administered. Each unit (participant) is in only one condition, and it is therefore a *between-participants* design. It is a true experimental design because individual test units are randomly assigned to the conditions.

Statistical rationale. We now explain the benefits of randomization. Suppose in the previous example G1 and G2 each consisted of 30 participants where each participant rated the brand on a five-point scale. We could then calculate the means for both groups and test if they are statistically different using a standard t-test or *analysis of variance* (ANOVA) (*see* ANALYSIS OF VARIANCE AND COVARIANCE) (assuming the observations have been measured at interval level) (*see* PRIMARY SCALES OF MEASUREMENT). If $O1$ has a statistically significant higher (or lower) mean than $O2$ this will be attributed to group 1 receiving treatment X. This is because all preexisting differences between individuals were controlled by randomly allocating them to groups 1 and 2. Adopting the conventional 5% level of significance means that chances of observing this difference based on mere chance (or actually, based on the combined effect of all unobserved "third variables") are less than 5%. That is, there is only 5% probability of drawing the conclusion that there is a difference if in "reality" there is no effect (this is called the *Type I error*). This percentage is normally deemed too low to accept the interpretation that the effects are caused by mere chance. Instead, the only other possible interpretation is accepted, which is that X caused the difference between $O1$ and $O2$.

If no significant difference is observed between $O1$ and $O2$ this may be either because there is really no effect of X, or the effect may have existed, but could not be detected. The latter means the design lacks sufficient "*power*" and the researcher runs the risk of committing a "Type II error": falsely concluding there is no effect. This can occur firstly when the sample sizes for the two groups are too small. Means observed for smaller samples display a larger variance and it is therefore more likely that differences of certain size occur even if there is no effect of X (*see* PROBABILITY SAMPLING). The researcher then cannot decide whether the observed difference is because of a real effect of X or because of a mere chance. A second possible reason for not detecting a true effect of X on O is the influence of other (unobserved) factors on the variation in units' scores for $O1$ and $O2$ (independently from X). These factors increase the variance in sample means for $O1$ and $O2$, making it difficult to detect any effect because of X.

G1 : X $O1$
G2 : $O2$

Figure 2 Two-group quasi-experimental design.

G1 (R): X $O1$
G2 (R): $O2$

Figure 3 Two-group posttest only experimental design.

These two reasons for a lack of power each have their own remedy : (i) the sample size can be increased, or (ii) the variation in the measures of $O1$ and $O2$ can be reduced. The effects of increasing the sample size are the same as for any random sampling: any doubling of precision requires a quadrupling of the sample size (or stated differently: precision increases with the square root of the sample size) (*see* PROBABILITY SAMPLING and STATISTICAL APPROACHES TO DETERMINING SAMPLE SIZES). Increasing the sample size is an obvious candidate for improving a design, but can be costly in terms of money and time. And if the researcher realizes the limited power only after the data have been analyzed (and no effects were found) it may be difficult if not impossible to replicate the exact experimental conditions to allow the collection of extra responses, especially given the requirement of random assignment. Merely topping up the numbers for one group is not sufficient; extra respondents should be randomly assigned as in the original experiment or one will end up with a quasi-experimental design.

Reduction in variance, which is the second method for enhancing power, can firstly be achieved by design. *Standardization* of test circumstances will reduce the variance in the dependent variable. Variance will further reduce if a more homogenous sample of participants is used, which is one reason why many consumer experiments in marketing are conducted on student samples (in addition to ease of access to such respondents by academic researchers). Instead of limiting the sample to a narrowly defined group of participants, it is also possible to include variables in the design that ensure that the sample can be treated as a set of homogenous subgroups. This is achieved by using *blocking factors* (discussed further below).

After the data have been collected, the variance can possibly be further reduced by means of statistical control. By including additional variables in the analysis that "pick up" or "explain" some of the variance in O the remaining variance to be explained reduces, thereby increasing the chances of detecting an effect of X, if it exists. In terms of data analysis this means that extra independent variables are included in the analysis as covariates using analysis of covariance (*ANCOVA*) or multiple regression methods. (*see* ANALYSIS OF VARIANCE AND COVARIANCE; MULTIPLE REGRESSION) Good covariates are variables that themselves are not affected by the treatment and that correlate substantially with the dependent variable (Tabachnick and Fidell, 2007). Typical candidate covariates are demographics such as age or gender, attitude measures — or prior observations on the dependent variable, which are called *pretest scores*. The next sections will discuss the use of pretest scores and the use of blocking factors.

Including pretest scores. *Pretests* are observations obtained for each test unit prior to receiving the treatment. By including such pretest scores in the analysis, the statistical power of the analysis can increase substantially and a same size difference between the posttest means can now become statistically significant. Such a design is called a "*pretest–posttest experimental design*." The basic form of this design is shown in Figure 4, for a case with one experimental and one control group. Note that pretest measures $O1$ and $O3$ are not used to directly assess the effects of X; they only act as control variables in the analysis. The actual analysis of a pretest–posttest design often involves recoding the dependent variables into a new variable representing the difference for each unit between the posttest and pretest scores.

In our example, we would for each participant calculate the difference between the pretest and posttest scores and use this value as the new dependent variable. This new measure will have less variance than the original scores because differences between participants' use of the measurement scale have cancelled out.

Although there are clear advantages to having a pretest measure, one should be careful that pretest scores are not used to create experimental groups as this may introduce a biasing regression to the mean effect (*see* VALIDITY IN EXPERIMENTATION). Another main concern is that exposure to the pretest measurement

G1 (R):	$O1$	X	$O2$
G2 (R):	$O3$		$O4$

Figure 4 Pretest–posttest experimental design.

may influence the participants such that the posttest scores no longer reflect the effect of X as intended but instead reflect the effect of "seeing the advertisement and having been asked to answer a pretest score." Such exposure to a pretest is a potential threat to the study's validity as it may sensitize respondents to the brand (*see* VALIDITY IN EXPERIMENTATION).

A design that allows detecting and separating these possible "*carry-over*" effects from the pure effects of X is the so-called *Solomon four-group design*, shown in Figure 5.

Comparison of the means of $O2$, $O4$, $O5$, and $O6$ will allow inferring the effect of X with and without prior exposure to the measurement instrument as follows:

- Effect of X without pretest exposure = $O5$–$O6$.
- Effect of X with pretest exposure = $O4$–$O2$.
- Effect of pretest only = $O4$–$O6$.
- Effect of X accounting for pretest scores = ($O2$–$O1$)–($O4$–$O3$).

It is sometimes possible to make more efficient use of available respondents and increase the sample sizes for different conditions by having respondents complete multiple conditions. If respondents are not completing one, but multiple treatment-measurement sequences, then more observations are obtained from the same respondent. However, this gain in efficiency comes with a cost, similar to introducing a pretest. The exposure to the first treatment may have carry-over effects and change the response generated by the second and following treatments. The use of such "*within-participants designs*" therefore, requires special precautions and they should be used only after consideration of these possible biasing effects, and how to account for them in the analysis. One compromise solution is to administer the second treatment and measurement, but include a sufficient number of respondents to allow conducting the main analysis on the basis of only the observations for the first treatment. This is shown in Figure 6. The first part of the design, with measures $O1$, $O3$, and $O5$, consists of a separate between-participants three-group design, and the total design is a mix between- and within-participants design. Analysis of within-participants designs typically involves so-called *repeated measures* analysis, which is a special type of MANOVA (*multivariate analysis of variance*) (Tabachnick and Fidell, 2007) (*see* REPEATED MEASURES ANOVA).

G1 (*R*) :	*O*1	*X*	*O*2	
G2 (*R*) :	*O*3		*O*4	
G3 (*R*) :		*X*	*O*5	
G4 (*R*) :			*O*6	

Figure 5 Solomon four-group design.

G1 (*R*):	*X*1	*O*1	*X*2	*O*2
G2 (*R*):	*X*2	*O*3	*X*1	*O*4
G3 (*R*):		*O*5		*O*6

Figure 6 Within-subjects design with control group.

Using blocking factors. The general discussion so far has assumed that we always aim to assign experimental units randomly to the various treatment conditions. There is however substantial efficiency to be gained from the incorporation of the so-called blocking factors. *Blocks* are existing groupings of units that are assumed to show systematic differences in their scores on the dependent variable. By completely or partly repeating the experimental treatments within each of the blocks, these differences can be taken into account in the analysis. This reduces the variance in the dependent variable, thereby providing more scope for the effect of the experimental treatment to be detected.

If we expect males and females to differ in how they perceive the brand which is the focus of our study, it would be wise to use this information when designing the study. To use gender as a blocking factor, we would set up the study such that an equal number of males and females are randomly assigned to the experimental and control conditions. This consists of a *randomized block design*: all experimental conditions (seeing the advertisement versus not seeing the advertisement) appear equally often within all blocks (genders).

However, it is not always feasible to administer all treatments in all blocks. Suppose for our study we intercept shoppers in a shopping center and wish to study multiple brands. Because the

questionnaire will become too long we decide to ask each respondent about only two brands controlling for order in which the brands are presented. In this case, an *incomplete block design* may be considered, which means that only a subset of all treatments (brands) is administered in each block (position in the questionnaire as first or second brand) while it is ensured that each of the treatments appears equally often with each other treatment. Incomplete blocks are only available for a specific number of treatments and blocks (Montgomery, 2004).

Yet another type of blocking design is the *Latin square design*. This design can be used if two variables are used as blocking factors. Each experimental condition is administered to only one combination of variables while ensuring that each combination of blocking variable and treatment occurs equally often. In our example, if we included three brands per questionnaire and were also concerned about sample differences based on day of the week (as different days attract different types of shoppers), this could involve that we vary the order of presentation of the brands and day of the week, handing out only one version on each day, as shown in Table 1. Such a design assumes that effects of day of the week and order of presentation are independent of each other. Like incomplete block designs, Latin squares exist for only particular combinations of factors (Montgomery, 2004).

Factorial designs. So far we have considered only two or three experimental groups, each receiving one or two different experimental treatments or comprising a separate control group. Each treatment however can be made up of a combination of factors. An *experimental factor* is an independent variable describing one facet of the experimental condition or stimulus. Factor *levels* are the discrete and fixed conditions for the factor to which units are randomly assigned.

Table 1 A Latin square design.

	Friday	*Saturday*	*Sunday*
Order 1	Brand A	Brand B	Brand C
Order 2	Brand B	Brand C	Brand A
Order 3	Brand C	Brand A	Brand B

Factorial designs have two or more factors that are "crossed" to observe all possible combinations of factor levels. The simplest factorial design is a 2×2 ("two by two") factorial design, comprising four experimental groups. Although statistically there is little difference between adding a blocking factor and adding another experimental factor, the important conceptual difference is that experimental units are randomly assigned to the conditions of the experimental factor while the conditions of the blocking factor comprise previously existing groups. With blocking, the researcher has no control over the allocation of units to blocking conditions and therefore, differences between block means are merely descriptive and cannot be used to infer causal effects of the block factors.

The main advantage of factorial designs is that they allow studying the effect of changes in two or more factors at the same time. They therefore allow for the exploration and testing of interaction effects. An *interaction effect* is any effect on the dependent variable that depends on a specific combination of values of independent variables. Interaction effects can be analyzed and statistically tested using ANOVA or multiple regression. To interpret an interaction effect it is usually advisable to plot the effects in a line graph (*see* ANALYSIS OF VARIANCE AND COVARIANCE).

Fractional factorial designs. Factorial designs are very useful and are the backbone of a lot of experimental research. They however have one main limitation: as the number of factors increases, the number of conditions to be observed increases dramatically. Each addition of one two-level factor results in a doubling of the number of conditions. This is easy to see: a 2×2 has 4 conditions, a $2 \times 2 \times 2$ has 8, and so on. A factorial design of 7 two-level factors comprises 2^7 or 128 different conditions, a full factorial of 15 two-level attributes consists of 2^{15} or 32 768 different treatment conditions. This number of conditions will often be too difficult or expensive to produce and/or to administer.

To overcome this limitation, fractions of factorial designs are often used. *Fractions* are balanced subsets from the full factorial design. They have the property that although not all

Table 2 2^3 full factorial design, alias 2^7 fractional design.

Factors	*A*	*B*	*C*	*A* × *B*	*A* × *C*	*B* × *C*	*A* × *B* × *C*
Treatments				=D	=E	=F	=G
1	1	1	1	1	1	1	1
2	1	1	−1	1	−1	−1	−1
3	1	−1	1	−1	1	−1	−1
4	1	−1	−1	−1	−1	1	1
5	−1	1	1	−1	−1	1	−1
6	−1	1	−1	−1	1	−1	1
7	−1	−1	1	1	−1	−1	1
8	−1	−1	−1	1	1	1	−1

possible interactions can be tested, at least the "main effects" of all factors can be observed and tested independently of each other. The smallest fractional factorial design with seven two-level factors comprises only eight treatment combinations; a fractional design for a study with 15 factors can be as small as only 16 treatment combinations.

To demonstrate the features of fractional factorial designs we follow up on our previous example of testing various advertisement formats. Suppose we wish to test how evaluations of our advertisement are influenced by various promotional and layout features, including mentioning of a price discount (yes or no), mentioning of product quality award (yes or no), order details (limited or full details), theme color (red or blue), advertisement size (small or large advertisement), font type (small or large font), and position on magazine page (top or bottom of magazine page). These are seven factors to consider. Each has two levels so the total number of possible advertisement configurations to test is 2^7, or 128 configurations. We can create a balanced subset of only eight configurations as follows.

We start with a three-factor full factorial design in eight treatment combinations (Table 2). Each factor is coded +1 and −1 to indicate the two possible levels. The first three columns dictate the level used in each treatment, the remaining columns are only used in the analysis to estimate all possible (two- and three-way) interaction effects. So, we have directions for creating the variations of discount, award, order details and will estimate their "main effects" from columns A, B, and C and use the remaining columns to estimate three two-way (or "first order") interactions (A × B, A × C, B × C), and one three-way (or "second order") interaction (A × B × C). The codes in these seven columns describe all possible differences between the eight conditions.

Fractional factorial designs utilize these interaction columns and "overlay," or *confound* them with new factors, assuming that none of the original interactions exist. This is the crucial assumption underlying fractional factorial designs: by assuming interactions to be nonexistent (or zero), they free up columns in the design table to be used as "main effects" columns for extra factors D to G, which similarly are assumed to not interact with any of the other factors.

Table 2 shows the full factorial design with three factors in eight treatments and demonstrates how the same design table can be used to generate a fractional factorial design with seven two-level factors. Across the eight conditions each factor level occurs equally often with each of the other factor levels. For example, there are four levels "1" for A and within these there are two occurring with level "1"for B and two with level "−1" for B; similarly there are two occurring with level "1" for C and two with level "−1" for C. This property can be tested by calculating the correlations between the seven columns in the design. All correlations should be zero. Designs as in Table 2 are called *"independent main effects plans"* or *"orthogonal designs."*

The principle can be extended to include more factors and to include factors with more

levels. For example, a main effects plan for 15 factors can be constructed that includes only 16 treatment combinations. A main effects plan of 12 three-level factors can be constructed that includes only 27 treatment combinations. More complex designs can be created in various ways, including designs for combinations of two and three or more levels per factor.

The above demonstrates the main principle behind the construction of a fractional factorial design. Not all fractional factorial designs assume all interactions to be zero and not all of them confound main effects with two-way interactions. So-called Resolution IV designs allow estimating main effects independently of such interactions, and other designs even allow estimating some interactions; however this comes always at the cost of requiring larger numbers of treatments (Montgomery, 2004; Hensher, Rose, and Greene, 2005).

The principle illustrated here also does not always result in the most efficient design. For a particular applied problem more optimal combinations may exist, depending on the number of factors and factor levels, the amount of prior information about variation in the dependent variable, the cost of creating different conditions, and the expected sizes of the effects. It is beyond the scope of this article to discuss these more technical issues except for noting that recently substantial progress has been made in developing efficient designs for studies with larger numbers of factors, especially in so-called choice experiments (Street, Burgess, and Louviere, 2005). Catalogs of design plans are readily available or design plans can be generated using software such as SAS. SPSS includes a routine for generating main effects designs.

Fractional factorial designs have been used, in particular, in two related methods for studying consumer preference and choice, called *conjoint analysis* and *choice experiments*. In *conjoint analysis*, respondents rate the attractiveness or purchase likelihood of hypothetical alternatives that are presented to them one by one (*see* CONJOINT ANALYSIS). Instead, in *choice experiments* respondents repeatedly make choices between different alternatives (Louviere, Hensher, and Swait, 2000; Hensher, Rose, and Greene, 2005). Alternatives can be products, destinations, travel options, sites, or any other relevant choice option. Alternatives are described by their main features, called *attributes*. These can be any feature like in case of travel options, mode of transport, destination names or types, travel cost, local attractions, accommodation types, and so on. Different combinations of attributes result in different hypothetical options. From the observed ratings and choices the researcher infers the relative importance of each feature, typically using regression-based methods. The approach uses experimental design methods, especially fractional factorial designs to create the different option descriptions.

CONCLUSION

The aim of most experimental research is to test causal relationships. Causal relationships are giving decision makers control and are a basis for deciding what actions to take to obtain desired outcomes. Experiments are powerful but require careful preparation. Also they can accommodate only a limited number of variables at any time. Experiments are therefore often conducted as a sequence of different tests. They are also typically only conducted after extensive exploratory research through interviews, surveys, or observational methods. The research literature obviously is another important source of information to allow the development of ideas in sufficient detail to make experiments worthwhile.

Finding a good design depends on the particular setting, what variables need to be controlled for, the opportunities for applying randomization, and the trade-off between internal validity, which requires homogenous conditions, and external validity, or the extent to which the test situation is representative of the target application situation. In addition to internal validity, which is about assessing causality, another main issue is construct validity: a treatment may have had a reliable effect and all alternative explanations have been ruled out, but often the question remains as to what is the actual "active ingredient" of the treatment, and does this active ingredient correspond with the constructs in our theories?

To conclude, experiments are the best method for testing causal relationships but only a limited

number of relationships can be tested in any applied setting. Understanding basic principles of experimental design is important even if one has no intention of conducting experiments as it will allow to better understand the limitations of other research approaches. Experimentation is not just a research approach; it is a mindset.

Bibliography

Cook, Th.D. and Campbell, D.T. (1979) *Quasi Experimentation: Design and Analysis Issues for Field Settings*, Houghton Mifflin, Boston.

Dunbar, R. (1995) *The Trouble with Science*, Faber and Faber, London.

Hensher, D., Rose, J., and Greene, W. (2005) *Applied Choice Analysis, A Primer*, Cambridge University Press, Cambridge.

Latour, B. (1987) *Science in Action: How to Follow Scientists and Engineers through Society*, Harvard University Press, Cambridge.

Louviere, J.J., Hensher, D., and Swait, J. (2000) *Stated Choice Methods*, Cambridge University Press, Cambridge.

Montgomery, D.C. (2004) *Design and Analysis of Experiments*, 6th edn, John Wiley & Sons, Inc., New York.

Shadish, W.R., Cook, Th.D., and Campbell, D.T. (2002) *Experimental and Quasi Experimental Designs for Generalized Causal Inference*, Houghton Mifflin, Boston.

Street, D.J., Burgess, L., and Louviere, J.J. (2005) Quick and easy choice sets: constructing optimal and nearly optimal stated choice experiments. *International Journal of Research in Marketing*, **22**, 459–470.

Tabachnick, B.G. and Fidell, L.S. (2007) *Using Multivariate Statistics*, 5th edn, Allyn and Bacon, Boston.

descriptive research

Vikas Mittal

Background

What is descriptive research? A survey of marketing-research textbooks shows there is no single definition of descriptive research. Some authors, for instance, argue that only cross-sectional research is descriptive research, separating it from longitudinal studies. Others exclude qualitative research from descriptive research. Contrary to this limited and limiting view, we believe that descriptive research refers to a broad category of marketing-research studies that are used to "describe" a state-of-affairs, phenomenon, and/or trend for the firm. As long as the primary goal of the research is to describe rather than assert true causality in an experimental sense, we view it as belonging under the broad umbrella of descriptive research.

Most descriptive studies in marketing research are typically cross sectional in nature, though they also include longitudinal studies, tracking studies, and, in some cases, qualitative studies. In summary, descriptive research conceptualized most broadly is all manner of marketing research that is used to portray any phenomenon of interest. From a statistical analysis perspective, part of the portrayal includes not only different *levels* of various constructs and variables (i.e., univariate analysis) but also the *association* among them (i.e., bivariate and multivariate analysis). Consider the following examples of descriptive research:

- To ascertain its brand image, a firm does a cross-sectional survey of customers of its own brands and competitors' brands on various dimensions. These results are then presented in the form of a perceptual map to the management.
- As part of its attempt to segment its customer base, a firm conducts a *cross-sectional* survey of its customers and potential customers. The survey data are merged with behavioral variables (e.g., television viewing habits). A cluster analysis is done to identify segments. The resulting segments are described based on brand perceptions and attitudes, behavioral intentions, and other descriptive variables.
- To ascertain factors driving sales, a business-to-business firm measures product attributes, perceptions of sales-force responsiveness, and pricing policy. A regression analysis is used to identify the key drivers of sales.
- Automotive firms use *syndicated research* sources such as J.D. Power and Associates to monitor customer satisfaction, TGW (things gone wrong), and reliability of their brand.
- Sales trends of consumer packaged goods are tracked by brand, category, and market from syndicated *longitudinal* data provided by firms such as Neilsen.
- To quickly ascertain the impact of changes made to its check-in system, a hospital conducts a focus group of recently admitted patients. Feedback from this *qualitative research* is used to tweak and refine the system for improvements.
- A survey of students at a university shows high levels of dissatisfaction with the student center. A series of focus groups are convened to identify specific areas of improvement.

All of these are examples of descriptive research. The primary focus of this article is on cross-sectional research, though we do touch on other forms of descriptive research as well.

Cross-Sectional Descriptive Research

Cross sectional, descriptive research is best described as a research study that aims to take a snapshot of the market place. In a cross-sectional marketing-research study, customers are surveyed at *one point in time*. On the basis of their responses to the survey, a set of conclusions that "describe" the phenomenon of interest can be drawn. There is a large variety of cross sectional, descriptive studies in marketing research (*see* SURVEY RESEARCH). Some important types of marketing-research studies using a cross-sectional design are described below:

1. *Buyer-behavior study*: In this type of a descriptive study, the firm wants to

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

describe various behaviors among its customers. These behaviors are not only those affecting the firms' brand directly, but also include any and every behavior that may be of interest (e.g., media-viewing habits, or behaviors involving competitor brands). Typical questions may include the following:

a. How do customers use our (and competitor) brands after buying them?
b. Where do customers go to purchase the brand they chose?
c. Are there differences in brand usage patterns between newly acquired customers versus loyal customers?
d. What are the different occasions for which customers use our brand (versus others)?

Then, different segments of customers are profiled and compared on the basis of a variety of behaviors. Notably, in many cases, these are self-reported behaviors.

2. *Brand perception study*: In this type of research, firms want to know how their brand is perceived by customers relative to other brands in the market place. Using a variety of rating scales, a single respondent may rate different brands on a variety of attributes. These can relate to brand attributes (e.g., brand personality) or product attributes, or a combination. The results of the study are visually depicted as a series of simple bar charts or, in some cases, a perceptual map created using a factor analysis or correspondence analysis. These positioning maps are used to identify the brand's position, vis-à-vis competitors, and among customers belonging to different segments.
3. *Customer segmentation study*: In this type of descriptive research, the goal is to obtain a random sample of all customers (current and potential) and measure variables, including psychographics, brand perceptions, demographics, usage, and so forth. Statistical techniques like cluster analysis can be used to divide the sample into segments. On the basis of the firm's objectives, the segments may be based on attitudes, perceptions, behaviors, or demographics (exclusively or some combination of these). Once the segments are profiled, they help the firm to understand the market, decide on the target segment(s), and devise a positioning strategy. Segmentation studies are most useful for firms trying to get a fine-grained description of the market. In many instances, firms may conduct a descriptive study that combines elements of all the three types of studies described. Thus, combining a brand perception and segmentation study can reveal the brand's positioning among a variety of segments. Later these segments can be profiled on the basis of specific behaviors to help with targeting.
4. *Customer satisfaction study*: Firms will typically survey a cross section of their customers to measure overall satisfaction, behavioral intentions, and attribute-performance perceptions. On the basis of a customer satisfaction study, the firm can gauge the level of satisfaction among customers, how they perceive the performance of the firm on various attributes, and their likelihood to repurchase from the firm. The firm can also ascertain attribute importance, or the extent to which an attribute affects the overall satisfaction. Results of these studies help firms to create importance-performance charts that guide resource allocation for customer satisfaction improvement.
5. *Lost-customer study*: Realizing the high cost of customer acquisition, many firms now want to understand reasons for which customers may have stopped doing business with them. In addition to measuring customer perceptions of their own brand, a lost-customer survey also measures perceptions of the brand to which customers switched. Critical incident information, for example, a short description of a particularly dissatisfying experience, is also obtained. Results can be used to create strategies to reclaim lost customers as well as to prevent customer loss. Some of these studies also help firms to identify service recovery strategies.
6. *Advertising evaluation study*: To evaluate its advertisements, firms often conduct advertising evaluation studies. Typically,

these are before and after studies, using a pseudo-experimental design. A random sample of customers is surveyed about firm perceptions before an ad is run in the market. After some time (typically, one to three weeks) of the ad being run, another random sample of customers is surveyed. The before and after samples can be compared to ascertain the difference in perceptions among consumers. Such studies lack internal validity and cannot be used to infer true causality. Yet, they provide firms with a quick reading of the ad's effectiveness.

7. *Employee attitude and satisfaction study*: Though mostly used by HR departments, experience shows that marketing-research departments typically take the lead in conducting employee satisfaction surveys. Furthermore, in service firms, these surveys are also used to assess the impact of employee satisfaction on customer satisfaction. Typically, a small sample of employees at each sub unit (e.g., bank branch, retail outlet, or dealership) fill out a survey that measures overall job satisfaction and its antecedents such as satisfaction with supervisor, pay, organizational processes, and so forth. In addition, these surveys can also be used to identify training and employee development needs. Sometimes, they are also used to measure perceptions of the corporate brand among employees.

As mentioned earlier, this is not an exhaustive list of descriptive studies using a cross-sectional survey. Rather, it illustrates the broad spectrum of goals that can be accomplished using cross-sectional market-research studies.

Survey methodology for cross-sectional studies. A variety of choices are available including, in-person surveys, telephone surveys, Internet-based surveys, and mail surveys. The choice of a survey methodology is based on several criteria including the following

- *Cost*: The total cost of conducting the survey. In-person surveys are more expensive than phone surveys, because the interviewer must travel to the respondent or vice versa. Typically, phone surveys are more expensive than mail and web-based surveys.
- *Speed*: In-person surveys can slow the data collection process because of travel time. Mail surveys, are somewhat faster, but telephone and Internet-based surveys are the fastest.
- *Versatility*: Mail surveys and Internet-based surveys allow respondents to evaluate visual stimuli (e.g., rate the different shapes or colors of a product or brand logo). On the other hand, telephone surveys and in-person surveys allow the interviewer to follow up and clarify the responses provided. Telephone surveys and Internet-based surveys also enable researchers to program complex skip patterns (respondents can be directed to different subsequent questions based on their answer to the present question), which is a weakness of mail surveys. Thus, each methodology is versatile in different ways.
- *Privacy*: Mail and Internet surveys afford the maximum privacy to respondents who can answer the questions in their homes. This is particularly important when studying sensitive topics or subject matters that elicit strong socially desirable responses. Telephone and in-person surveys afford relatively less privacy.
- *Cultural sensitivity*: When conducting surveys, attention should be paid to the local context. For instance, India and Mexico have relatively underdeveloped telephone service, and the mail is unreliable. Thus, it is more useful to conduct in-person surveys.

Many other factors besides these should be considered when making the choice of a survey methodology. More and more, companies are opting to use Internet-based surveys as they provide the speed of a telephone survey as well as the visual elements of a mail survey. However, Internet surveys rely on email lists that may be substantially biased on the basis of computer ownership and/or access. Still, with increasing Internet access this may be less of an issue.

Sample design for cross-sectional studies. The external validity, that is, the extent to which

results can be generalized to a particular population, of a cross-sectional survey critically depends on the quality of the sample chosen. Theoretically, the best sample is one where the units included in the sample are randomly chosen from the population of interest. However, in reality, all units of the population are not available for potential inclusion, as the random selection is done from a sampling frame, or a list of all available population units (*see* SAMPLING TECHNIQUES; PROBABILITY SAMPLING). For instance, a firm wanting to create a random sample of its customers may choose to select customers from its email database. To the extent that some customers simply do not have emails, they will be underrepresented in the sampling frame causing the final sample to be *underadjusted* for those of the firm's customers who do not have emails. It may also be that some of the people whose email addresses are in the list are no longer with the firm. In this situation, the sampling frame will be *overadjusted* for noncustomers. Thus, because of underadjustment of the sampling frame some units who should be part of the random sample are not, and because of overadjustment some units who should not be part of the random sample are included. If these two subsets are systematically different, then the ultimate random sample can be biased. For instance, consider a local bank in Houston. Many of its ultrawealthy clients do not provide an email address, and many student clients whose address the bank has have migrated to other cities. Given this, any random sample drawn from its email list will be biased in representing the firms "current customer base." Thus, when creating a random sample, biases because of the composition of the sampling frame should be explicitly noted.

The marketing-research manager should also carefully consider the decision about the desired sample size. Clearly, a larger sample is desirable as it reduces the sampling error, that is, the margin of error (at a given level of confidence, say 95%) associated with the study. However, for a given budget, increasing the sample size also increases the nonsampling error because of nonresponse bias and administrative errors. In many situations, nonrespondents are systematically different from those responding to the survey. Thus, even with a large sample size, a high nonresponse bias can seriously undermine the representativeness of a study. The best strategy is to have a smaller sample size, but use the available budget to increase the response rate. This will ensure that all the units randomly selected from the sampling frame for study inclusion are, in fact, part of the study. In many cases, the length of the survey and the design of a survey (e.g., uncluttered survey with lots of white space) can significantly impact the response rate. For instance, longer surveys not only decrease response rate (thereby requiring a larger sample to obtain a prespecified number of completes) but also necessitate higher respondent incentives.

Constructs measured in cross-sectional descriptive studies. The constructs measured in a descriptive survey consists mainly of self-reports on a variety of domains including customer perceptions, attitudes, intentions, behaviors, and background variables. Broadly speaking, they may be classified in the following categories (*see* QUESTIONNAIRE DESIGN).

- *Brand perceptions and attitudes*: Customers have perceptions of and attitude toward a brand or a firm. For instance, scholars have shown that a positive brand attitude helps a firm to insulate its sales in case of a negative crisis. In a typical cross-sectional study, brand perceptions and attitudes can be measured using Likert or Semantic differential scales. Examples are as follows:
  - How would you rate Money Bank on the following dimensions:
    - Safe–Unsafe; Friendly Staff–Unfriendly Staff
  - Rate your agreement with each statement:
    - My assets are safe at Money Bank (1 = strongly disagree, 10 = strongly agree)
    - The staff at Money Bank is friendly and courteous (1 = strongly disagree, 10 = strongly agree)
- *Behavioral intentions*: Scales with end points such as "extremely likely" and "extremely unlikely" are used to measure various intentions such as intent to purchase a brand,

intent to repurchase a brand, intent to engage in word of mouth, and intent to recommend. Typically, these intention variables are used as surrogates of actual behavior. However, though positively correlated, the strength of the relationship between attitudes and behaviors varies widely. Thus, the presence of high intention scores should not be interpreted as correspondingly high behavior among customers.

- *Stated behavior*: Surveys also ask consumers to retrospectively state a variety of behaviors. For instance:
  - In the past six months, how often have you visited a Greenery grocery store?
  - What percentage of your disposable assets is invested with firm X and what percentage with firm Y?
  - In the past months, to how many people have you recommended Brand Y?

  Despite their limitation of one-to-one correspondence with actual behavior, stated behaviors are used extensively to understand customer behaviors.
- *Demographics*: Customers typically fill out a background section measuring demographics such as gender, race, income, marital status, and so forth. These are used as classification variables in the analysis.

The above is a list of the major categories of variables measured in cross-sectional descriptive studies. Note that the specific variables measured depend on the needs of the firm.

Analysis of cross-sectional descriptive data. Typically, descriptive cross-sectional studies are analyzed using a variety of analyses. An overview is provided in this section.

Univariate analysis. In univariate analysis, the goal is to describe each variable one at a time. This can be done using either frequency counts (or percentages) or averages. Consider the variable, "My assets are safe at Money Bank," which is rated on a 10-point scale (1 = strongly disagree, 10 = strongly agree). The analyst may calculate the average score on this variable, say 7.5. Using percentages, the analyst may report the top-2 box score (percentage of people giving a rating of 10/9), say 82%. Depending on the firm's needs, the analyst may report top box scores (percentage of people giving a rating of 10), or top-3 box score (percentage of people giving ratings of 10/9/8), or any top-X box score. Typically, the average score and the top-X box score are positively correlated though the magnitude of the correlation depends on the distribution of the scores across respondents.

In a marketing-research report, the average scores or top-X box scores are reported as simple tables or bar charts. To this end, the analysis can usually be conducted using a spreadsheet or a basic statistical analysis package. Despite the simplicity, the analysis is useful in many respects. It provides a quick snapshot of all variables in the survey. In many marketing-research situations, this is all that a manager needs to draw conclusions or conduct basic hypothesis testing. For instance, on the basis of the observed top-2 box score of 82%, the manager can statistically test the hypothesis that in the population of interest the top-2 box score is 75%.

Bivariate analysis using cross tabulations. The goal of bivariate analysis is to analyze two variables at a time. In most marketing-research studies this is done using two-way cross tabulations, known as *tabs* (*see* CROSS-TABULATION). For any given study, the client and the marketing-research firm decide on the format of the tabs. The most important decision is to agree on the banner and the stub. The banner of a cross tab represents the column variables and typically consist of segmentation variable(s) by which the data is broken down. The stubs are the rows, and consist of the variables that are expected to vary by the segments, or by the banner.

For the variable, "My assets are safe at Money Bank," which is rated on a 10-point scale (1 = strongly disagree, 10 = strongly agree), suppose the desired banners are gender, income, and customer status each having the following categories:

- Gender (male/female)
- Income (high, medium, low)
- Customer status (current customer/non-customer).

It should be obvious that these three banner variables lead to a total of seven banner points

Table 1 A cross table showing banner points and stubs.

	Gender		*Income*			*Customer Status*	
	Male	*Female*	*High*	*Medium*	*Low*	*Current*	*Noncustomer*
Top-2 box score (%)	80	75	70	90	X	90	Y

(2 + 3 + 2). Using these banner points, the cross tables describe the percentage of people who fall in each category of the questions analyzed in the stub. For instance, the top-2 box score for Gender may be displayed as in Table 1.

The rows can be as detailed or as aggregated as desired. For instance, the cross tabulation can show the number and percentage of people giving each rating on the 10-point scale, for a total of 10 stubs. Additionally, the cross tables can also report the average, and the top-X box scores. Thus, for a 10-point scale-item, one can have more than 10 stubs.

When preparing a market-research report, it is customary to display these data in the form of bar charts, with different colored bars representing each banner-point. Such bar charts are simple to create and visually easy to understand. This simplicity, as explained earlier, and the associated speed, is an important reason for the widespread popularity of cross tabulations. From a statistical analysis perspective, cross tabulations allow people to test hypotheses about the independence of two categorical variables using the chi-square test. They can also be used to compare the means (z-test or t-test for means) or percentages (z-test for proportions) based on two groups. Such hypothesis testing enables the manager to compare how two groups are different with respect to a focal variable.

Multivariate analysis of cross-sectional data. In many cases, multivariate techniques are used to analyze cross-sectional data for descriptive studies. A complete discussion is beyond the scope of this article, though a brief overview with examples is provided.

- *Perceptual mapping*: One set of multivariate techniques such as factor analysis and correspondence analysis is simply used to characterize respondent perceptions in a visually meaningful way. Thus, perceptual maps are created from ratings of multiple brands provided by the customers. Perceptual maps provide a visual representation of various attributes and brands to identify the underlying patterns in their interdependence.
- *Cluster analysis*: This analysis classifies the respondents in the dataset based on their similarity on some prespecified criteria (e.g., attribute ratings) (*see* CLUSTER ANALYSIS). The most common clustering technique used to accomplish this in marketing research is called *k-means clustering*, though advanced techniques such as latent segmentation are also employed.
- *Multiple regression analysis*: In this technique, a set of *independent* variables is used to predict a *dependent* variable (*see* MULTIPLE REGRESSION). For instance, overall brand attitude is predicted on the basis of specific brand perceptions of "exciting," "dependable," and "novel." Or, customer-share of wallet may be predicted on the basis of their satisfaction with the firm's service, brand commitment, and household income. It should be noted that even though there is a dependent variable in regression analysis, the analysis makes no presumption about causality – that is, the independent variables are simply associated with the dependent variable, rather than causing it in an experimental sense.

Limitations of cross-sectional analysis. A key limitation of cross-sectional research is its inability to provide causal conclusions. In other words, all variables are simultaneously measured in a single point in time, and thus, temporal sequencing – a precondition for establishing causality – is missing. Nonetheless, in most cases, managers use their mental models to impose causal ordering on the variables measured in the survey. For instance, based on

their experience, and causal research conducted in the past, the managers may have a mental model of causality that an increase in promotion through coupons causes unit sales to increase. From this mental model, a manager may feel comfortable interpreting the cross-sectional results of a correlation between the amount of promotion and unit sales in a "causal" sense. Clearly, the magnitude of the association may still be inflated owing to common methods bias inherent in cross-section research.

Another limitation of cross-sectional research is that it does not allow one to observe changes within a unit. Thus, a person's perceptions of a brand may change over time. However, if we happen to measure it at a time when they are high (e.g., right after an advertising campaign) and we mistakenly conclude such perceptions are stable, we may describe them as being "high" even after some time has lapsed. In reality, the perceptions may have declined after a period of time, leading to an incorrect conclusion.

Conclusion. It should be apparent to the reader that descriptive research using a cross-sectional design incorporates many elements about which decisions need to be made. Each decision has a critical impact on the survey outcome, and the nature and type of conclusions that can be drawn from the study. Thus, it is advisable to obtain guidance from marketing-research experts when designing the study, particularly in regard to elements like sample design, survey design, and analysis plan. Some sources for guidance include Alreck and Settle (2004), Parasuraman, Grewal, and Krishnan (2007), and Salant and Dillman (1994).

Longitudinal Descriptive Research

In contrast to a cross-sectional study, which is best described as a *snapshot*, a longitudinal study is like a *movie*. In a longitudinal study, data from the same entity is gathered over multiple periods of time. For instance, the same customers may be interviewed each month to obtain data on what groceries they purchased each month and their spending patterns. An automotive buyer may be surveyed within two weeks of buying a new car, after six months of ownership, and then again after three years of ownership. By interviewing the same entity over time, the researcher can draw conclusions about changes in the *levels* of variables of interest. The ability to observe changes in the variable of interest is a key benefit of longitudinal research as compared to a cross-sectional study.

In most research settings, interviewing the same-people by tracking them over time can be expensive and time consuming. Even if the firm has the necessary financial resources and time to devote to a longitudinal study, several problems may beset such a study. First, respondents may drop out of the study for reasons that are beyond the control of the firm. These may range from mortality to relocation to simply fatigue from repeated participation. Second, there may be sensitization and measurement reactivity. In other words, when the respondent is filling out the same survey again and again, the mere act of doing so can sensitize respondents and change their responses. Using a within-subjects analysis or a repeated-measures analysis can mitigate this concern to some extent.

Relative to cross-sectional studies, longitudinal studies are less often used in the marketing-research industry. This is not a limitation of longitudinal studies, but may merely reflect the desire for cost-effective ways of addressing marketing-research needs. It should be noted that longitudinal studies lend themselves to statistical analysis that enables stronger causal conclusions and allows researchers to control for unobserved heterogeneity.

Tracking Studies

A tracking study is different from a longitudinal study described above. In a longitudinal study, the same entity (e.g., customer) is surveyed or observed over different time periods. Thus, in a longitudinal study, the same customer would provide data over several time periods. While such a longitudinal study could well be a tracking study, the typical tracking study in marketing research is based on aggregating data from cross-sectional studies over time. For example, a firm may conduct a cross-sectional survey of customers who visit its store in Ohio every quarter to gauge brand perceptions. A similar survey, for the sake of example, is also done

on behalf of its store in Texas. In other words, a random sample of customers in Ohio and Texas is interviewed every quarter. Suppose for the Ohio store, the scores for seven consecutive quarters are 5.4, 6.5, 7.1, 7.4, 7.9, 7.8, 7.9 (perceptions measured on a 10-point scale). Similarly, suppose the scores for the Texas store are 9.5, 9.5, 9.1, 8.9, 7.5, 6.1, 6.4. These scores, which comprise the data for the tracking study, have several properties:

- They are based on cross-sectional surveys of customers at each store. Thus, the tracking study is not a longitudinal study at the customer level: every quarter a new cross-sectional sample of customers is surveyed. In other words, it is not the same customer who is followed up over seven quarters. Rather, results from seven separate cross-sectional surveys are averaged and used to monitor a trend over time.
- The average scores at the store level are "tracked" over time, and treated as having longitudinal properties. Thus, by aggregating (e.g., averaging) cross-sectional data from individuals over time, a longitudinal time series can be created at a more aggregate level of analysis (i.e., the store) to which the individual contributing the cross-sectional data belongs.
- Several examples of such tracking database can be found. The American Customer Satisfaction Index creates a satisfaction score for each firm on the basis of a cross-sectional survey. The scores for each firm are tracked over time, and treated as having longitudinal properties. Consider the Consumer Confidence Index. It is based on a cross-sectional sample of consumers every period. However, aggregated for the entire United States, the index is treated as a longitudinal variable. For an individual firm, examples of variables that may be tracked over time include
  - top-X box scores on variables like satisfaction, brand perceptions, and repurchase intentions;
  - average ratings on attitude, perception, and intention variables;
  - percentage of people engaging in a certain behavior over time (e.g., percentage of people complaining, or calling customer service);
  - indexes such as net promoter score.

Qualitative Descriptive Research

Descriptive research can also be classified on the basis of whether it is qualitative or quantitative research. The preceding sections of this article have focused on quantitative descriptive research. A key feature of quantitative research is that the data collected for such research is typically collected in a structured manner from a relatively large sample of customers. Qualitative research, on the other hand, primarily relies on unstructured data gathering that is designed to gain insight into the "why" and "how" of various phenomenon.

In most marketing-research applications, qualitative research takes the form of focus groups or in-depth interviews conducted with key informants chosen on the basis of the researcher's judgment. Different from observational research (e.g., ethnography), focus groups and in-depth interviews rely on people's ability to verbalize their thoughts and feelings about the topic at hand. Because this is a relatively difficult task, a trained moderator or interviewer guides the participant using questions, probes, and other facilitating techniques to elicit information. For example:

- A health maintenance organization (HMO) conducted a survey and found that 75% of its members wanted the physician to "treat me with respect." However, it was not clear to the HMO what specific actions physicians could take to convey respect. To this end, a series of focus groups were conducted to elicit specific encounters that were perceived as respectful and/or disrespectful and why.
- A Montessori school wanted to conduct a survey and find out how people view Montessori education as different from a traditional education. However, there was no battery of items that could be used in the survey. Qualitative research was used to understand people's perceptions of and attitude toward a Montessori education, and from there survey items were developed.

The first example should, at least, illustrate that qualitative research is not always a precursor to quantitative research. It is true that in a majority of the cases qualitative research is used to help researchers understand the phenomenon in question, and aid in designing the quantitative survey. In some cases, however, qualitative research has been finding its use as a back-end clarification and amplification tool designed to help researchers interpret findings from quantitative studies. In many cases, we find that supplementing quantitative research with qualitative research can improve the descriptive power and insights manifold.

Conclusion

It should be evident to the reader that the term *descriptive research* is descriptively simple but deceptively simplistic. It applies to a very broad range of studies, and descriptive studies can be classified in many categories. For instance, they may be classified as qualitative or quantitative descriptive research. They can be classified as cross-sectional or longitudinal descriptive research. They may be classified on the basis of the substantive nature of the study into buyer-behavior studies, brand perception studies, customer segmentation studies, or customer satisfaction studies. Descriptive research also includes syndicated studies. Among all the quantitative studies, cross-sectional surveys, and the resulting tracking study, remain the most widely used descriptive research.

Looking to the future, we expect wider use of longitudinal studies. Changes in information technology have allowed firms to monitor and track consumer behaviors in ways that are relatively unobtrusive. They also allow firms to administer surveys in ways that are more cost effective. Thus, we should see a gradual uptake in the use of longitudinal studies. We also expect more use of multivariate data analysis techniques. Software packages such as SPSS and SAS have allowed users to conduct these analyses with relative ease. In using these software packages, dangers associated with "little knowledge" should be acknowledged. We have seen many instances where multivariate analysis was conducted simply because it could be done, or cases where knowledge of the software was confused with a mastery of the statistical technique. Finally, we expect firms to be using various forms of descriptive research in a more integrative manner. The divide between qualitative and quantitative research, especially, should be bridged to gain richer insights that also have high levels of external validity.

Bibliography

Alreck, P.L. and Settle, R.B. (2004) *The Survey Research Handbook*, McGraw Hill-Irwin.

Parasuraman, A., Grewal, D., and Krishnan, R. (2007) *Marketing Research*, 2nd edn, Houghton Mifflin Company.

Salant, P. and Dillman, D.A. (1994) *How to Conduct Your Own Survey*, John Wiley & Sons, Inc.

ethics in marketing research

Marcus J. Schmidt

In recent decades there has been a growing trend to base strategic marketing decisions on market research data. The periodic survey of marketing research sponsored by the American Marketing Association (AMA) investigates the vast application of marketing research techniques in American businesses. When comparing results of 1988 and 1997, it turns out that the use of research methods activities has increased within 28 out of 34 categories (test marketing, consumer satisfaction studies, etc.). A study carried out by a Danish market research agency (Jysk Analyse) several times between 1984 and 2007 shows a similar trend. While no definite estimates exist, it appears that companies' financial resources invested in marketing research on average have gone up during the last decades. The growing use of marketing research studies by companies and public authorities has resulted in an increased focus on ethics in marketing research.

Definition

Marketing research ethics derive from a marketing researcher's relations with the parties in the research process, including respondents, interviewers, the research firm, clients, and the general public. A researcher has the responsibility to treat respondents fairly by being truthful to them about the nature and purpose of the research. At the same time, he or she has the responsibility to gather accurate and reliable data for the client. Whenever the fulfillment of these responsibilities creates a conflict, a research ethics problem arises. Unethical behavior is a form of misconduct that appears if those involved in the research process violate proper marketing research principles (Akaah, 1990).

Figure 1 provides an overview of the key elements involved in a marketing research process.

All elements may cause ethical problems or dilemmas:

Respondents now and then provide answers knowing that they are incorrect. For instance, the person may underreport his or her consumption of alcohol or overreport purchase of fruits because drinking too much is regarded unhealthy and thus socially undesirable, whereas eating fruits is regarded healthy and therefore considered socially acceptable. The issue of *social desirability* is addressed in another article in this encyclopedia (*see* SOCIAL DESIRABILITY BIAS). While a respondent's refusal to provide a correct answer cannot be regarded unethical, it may be inappropriate to deliver results to the client or to publish findings without mentioning the possible bias.

The *questionnaire* may contain questions that do not interest the respondent. Thus he or she will be less involved in filling in the answers. Also, the questionnaire may address sensitive topics: questions about sexual behavior, usage of drugs, religious beliefs, political preferences, and so on. It is known from research that sensitive questions often result in invalid answers.

Questions may be formulated in such a way that responses serve the interests of the client.

A private company obtained a license from the government to exploit a country's subsoil for 99 years. The company was allowed to drill for oil and sell it by paying an annual fee to the government. Once large oilfields were found and the company's profits began to swell, some politicians and experts began arguing that the contract was too attractive to the company and that the government needed to renegotiate it.

The editors of newspaper A were in favor of a renegotiation. They had an agency asking a random sample of respondents the following question:

> The government has given private company X a monopoly on drilling oil in country Y's subsoil. Members of parliament now discuss the possibility of an expropriation – a statutory confiscation entailing economic compensation. If the proposal is adopted other companies – private as well as state-owned – are allowed to drill for oil along with company X. Are you in favor of or do you oppose this?

Soon afterward, newspaper B – being against the renegotiation – had another agency ask *its* random sample of respondents:

> The parliament must soon decide upon a bill implying that the government breaks its

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

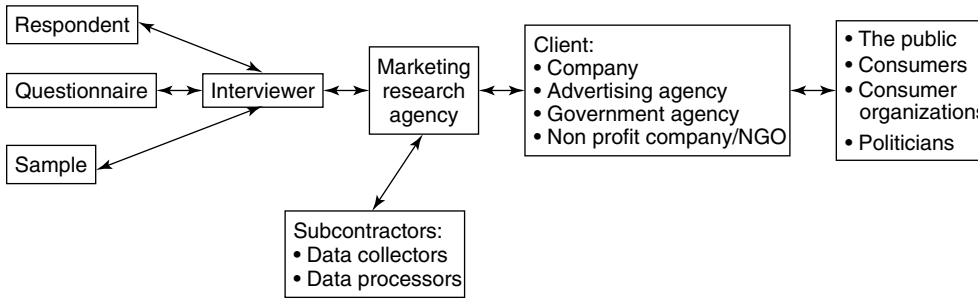


Figure 1 Potential ethical issues.

> agreement with company X about drilling. What do you think: Should the government comply with the agreement or do you find it acceptable that the government breaks the agreement? (OK/Not OK for the government to break agreement).

It is of little surprise that a majority answered "in favor" to the first question while a majority answered "not OK" to the latter question. Notice that in the present example – especially in the first case – it is not the question as such that causes a problem. The problem is that respondents are given "a helping hand" with regard to how the critical terms included in the text are to be understood.

In the former case, we are informed that an expropriation is in accordance with the law and that an economic compensation is being paid. Moreover, a monopoly business is transformed into a business with competition. In the second case, it is stated that passing the bill implies that the government breaks an agreement. For many people, a monopoly is associated with negative values (high prices, little product development, low service, big profit), whereas competition is regarded positively because of the opposite arguments. At the same time, it is broadly expected of a democratically elected government to comply with laws and agreements it has signed.

Both questions violate an ethical principle of avoiding leading formulations. (Both surveys were conducted by subsidiaries of renowned international marketing research agencies). An interesting observation: It turned out that the CEO of the oil-drilling company was a member of the board of directors of newspaper B. Also the top manager of the agency doing the study for newspaper B was a close personal friend of the newspaper's editors.

A careful researcher should spend sufficient time designing a questionnaire (see also Sudman, Bradburn, and Wansink (2004)).

The *sampling frame* may favor the interests of the client: Two competing trade magazines were eager to inform readers that *their* magazine was the preferred source of information among business managers and thus the most cost-effective medium for placing advertisements. Magazine A's core readers were known to be mostly small companies while magazine B's core readers were primarily middle-sized and big companies. Magazine A asked its market research agency to base the sample frame on small companies, whereas magazine B asked its agency to concentrate on middle-sized and big companies. It turned out that the overlap of sampling frame was as low as 5%. Consequently, based on their respective samples, both magazines could announce publicly that they were the preferred magazine among business decision makers.

The researcher should aim at selecting the most valid and reliable sample given the research problem at hand and not one that best serves his or her own interest.

The *interviewer* may bias results in several ways (see Table 1).

Subcontractors. Use of data collectors and data processors incurs additional problems. Each supplementary external element that is added to the research process enhances the risk of bias and of ethical problems. Data collectors may miscomprehend the sampling frame or they may be unable to fulfill sample quotas (i.e., that at

Table 1 Systematic interviewer bias.

In 1994, a German interviewer, Heiner Dorroch, published a biographical book, based on his life as a very "productive" interviewer for several big market research companies (Dorroch, 1994). According to Dorroch, interviewers are badly paid and therefore they are pressurized into faking interviews in different ways: often questionnaires used for personal interviews contain 500 questions or more (!)

Being aware of the challenges that long interviews cause for interviewers, research agencies have developed quality control procedures. For instance, the agency phones a sample of respondents (i.e., 10% of the sample) and asks them if they have indeed been visited by the specified interviewer on a given day. Since Dorroch had himself been involved in designing control checking procedures he knew exactly how to circumvent them. He comprehended that controllers would ask if he had been visiting a respondent's address. Also, he realized that they would confront the respondent with a few control questions: Did the interviewer ask the respondent to rate advertisements? Was the respondent exposed to a blind test of three cheeses, and so on. However, what the controllers did not check was *the length* of the interview and the number of questions asked. So, Dorroch was able to cut down the interview to about 20% of the intended length. Afterward he would fill in the missing answers by himself. Provided that a respondent favored ecological views during introductory screening questions, Dorroch assumed that that the person would purchase more ecological products and that he would possess more critical views regarding a company's social responsibilities than the average respondent and so on.

When interviewing managers, he frequently experienced that the interview was interrupted because the manager was too busy to complete the interview. In such cases, it became common to have another manager continue and complete the same questionnaire. Dorroch managed to work for several marketing research companies at the same time. So, during an interview, he would confront the respondent with questionnaires for several agencies. When he was fired at one agency he continued working for the same agency as a subcontractor for another interviewer. Dorroch also reports about a survey for a construction company where the payment of interviewers was directly associated with the number of confirmative responses: the more respondents that reported to be unsatisfied with their present apartment, the higher the payment to the interviewer and the more orders the client could expect to get from government authorities. As a consequence, the estimated need for reconstruction of apartments was significantly overestimated.

According to Dorroch, it is a major problem that interviewers are paid per completed interview. And because of the low earning per interview, only a very productive interviewer is paid an amount of money that compares to that of an average industrial worker. Dorroch claims that the procedures described in his book are more or less standard practice among interviewers.

least 80% of households own a DVD-recorder), while data processors may compromise on data quality checks or may use inappropriate weighting procedures.

The marketing research agency[1]. Problems attributed to employing a subcontractor also apply to the agency if the activity is carried out internally rather than being outsourced or offshored.

Sometimes, marketing research companies have been accused of reporting unrealistic return rates. In a pan-European study conducted by an international agency, it turned out that response rates were around 90% in a South European country while it was only 40–50% in Scandinavian countries – although the contact method was the same (telephone). Did the subsidiary with the 90% rate only count interrupted interviews as nonresponse but not missed calls and refusals? In the present case, the subsidiary was not at all cooperative regarding specifics about the details.

Owing to different reasons, agencies frequently refuse to unravel details about a research study it has conducted:

1. Doing so might give competitors access to confidential information about the client.
2. The research model is unique and is not to be disclosed to outsiders.
3. Giving access to details violates the anonymity of respondents.

Assume that an independent researcher learns about a study, say, at a seminar or in the trade press. He or she wonders about the results and finds them to be strange, confusing, or not in accordance with what he or she would expect. Therefore, he or she would like to check the validity of the research study. However, owing to one or several of the above reasons, the agency refuses to provide the necessary details about the study. In this case, it might be regarded unethical *not* to provide the researcher with technical details. On the other hand, providing details to the researcher may *also* be regarded as unethical: First, handing over the relevant information to the researcher may be regarded unfair to the client. Second, it may provide outsiders and thus competitors with access to the research agency's toolbox ("magic formula" or "secret recipe") and third, respondents may feel that the agency violates their integrity and anonymity. To sum up, this is a typical example of a clash of conflict between: (i) transparency of research and (ii) confidentiality.

Many marketing research agencies today use branding for protecting their product offerings. For instance, a consumer products company wanting to analyze the loyalty of its customer base can choose from among *Loyalty Plus*® from GfK, *Conversion Model*™ from TNS Gallup, and *BASES*® from Nielsen – just to mention one product from three different research suppliers. The common denominator of such branded products is that they are proprietary services, which means that the supplier refuses to disclose exactly how the research is conducted, which buyer-probability algorithm is employed, how the model produces repeat purchase rates, and so on. The underlying rationale of branded products (also known as *black box branding*) is to charge a premium price for the product service. For the client, such branding has the disadvantage that it becomes very difficult to compare the validity and quality of alternative approaches and figure out which one serves his or her interest best, given the specified price. While confidentiality with regard to branded models cannot be perceived as unethical, a common effort by the industry to establish a set of common ethical standards or principles for, say, conducting studies of consumer loyalty definitely would be helpful to clients.

The client. According to most textbooks, the client is a private company. But sometimes, the agency's contact is a subcontractor of the client company, that is, an advertising agency. In other cases, the client may be a government agency, a nongovernmental organization NGO or even a political party. The way a client treats results obtained from the research agency involves a lot of latent ethical problems. The client may have a hidden agenda and provide the agency with "appropriate" input with regard to the questionnaire, sample, which variables to use (and which not), and which research finding to report.

Why should a client want to influence the research process? In a private firm, there may be opposing interests within the management: The financial manager wants to cut costs and does not wish to spend money on a new product, whereas the marketing manager thinks that a new product is needed because the company is losing market share to its competitor. The marketing manager commissions an agency to carry out a research study. Clearly, he or she is interested in results that are supportive to his or her agenda: should the research design begin with a screening interview where only consumers who show positive interest toward the new-product concept are asked to participate in the focus group or in the market test? No doubt this would bias results yielding higher acceptance rates and purchase probabilities compared to a true random sample.

While the marketing manager's behavior represents an act of misconduct in a company context, it does not directly affect companies or persons outside the company.

But, suppose that the marketing manager willingly misrepresents or manipulates a customer satisfaction study, say, by merging the response category "don't know" with those being satisfied or by ignoring that the nonresponse category is overrepresented by former customers. Next presume that the satisfaction figures are used in

an advertisement ("88% of our customers are satisfied" – source: Agency XYZ). This would be an example of unethical behavior since it misrepresents the research carried out by the agency. Moreover, it deceives consumers by exposing them to information about the company that is not true (too optimistic).

It happens that the company's management refuses to use or publish findings either because it does not believe in the findings or because it finds them politically incorrect. A businessman once hired a group of graduate students to carry out a study about the need for a new trotting track in the region where he lived. He was himself an enthusiastic believer in the trotting-track project. But since he needed to convince the municipality, supportive research findings were required for publication in the local press. After having conducted interviews with potential customers and experts, the students concluded that the probability that the trotting track would ever become profitable was almost zero. There were several successful competing tracks at a driving distance of within one hour. The competitors had big resources for marketing campaigns and for hiring the best jockeys and horses. Moreover, loyalty among visitors of the existing tracks was generally high: Very few of those being interviewed at competing tracks indicated an interest in shifting to the new intended track. The report obtained a top grade at the business school. However, when the businessman read the report he became very angry and even called the business school's president complaining about the "negative" and "unproductive" report. The businessman decided not to publish any of the findings. Instead, he continued lobbying for the track (It was never built, though).

In the present case, we must regard the businessman's behavior as unethical.

In the following section, we will look at some more relevant dilemmas in the marketing research process.

Respondent – Agency Dilemmas

Some examples causing ethical problems are as follows:

1. An agency sells the respondents' names and demographic data to another company without their approval.
2. Respondents are contacted by a company whose purpose is to sell products. But the company disguises the true intention as an interview belonging to a survey (the contact form is sometimes called *sugging*).
3. Some persons choose not to have their phone number listed anywhere (secret numbers). Nevertheless, the number is randomly selected by a computer and the person is called by a research agency.
4. Respondents (i.e., employees of a supermarket owned by the client) are contacted by the agency's interviewers disguised as customers. The purpose is to assess how the respondents react to requests from customers. The procedure is known as *mystery shopping*.
5. An agency collects unobtrusive data from consumers who do not know that their behavior is being recorded – for example, assembling trash from households, carrying out interviews while using a hidden tape recorder, observing shoppers through a one-way mirror or letting respondents read a magazine that has previously been contaminated by a chemical liquid. An overview of *observation methods* is provided in elsewhere in this encyclopedia (*see* OBSERVATION METHODS).

The first two examples must be regarded unethical while the last three have been used by market researchers for many years.

Calling a person, who has done something actively for *not* being listed, causes an ethical dilemma. The person wishes only to be called by family and friends that have been provided the number but not by others like an interviewer or seller.

Several European consumer organizations have complained about calling unlisted numbers. Indeed, a market research company could contact the telephone company and get a file with the listed numbers. The sorted file could then be merged with a sorted file of the randomized numbers and only when there are duplicated numbers in a row the number may be called. Since all telephone companies must

be contacted, the procedure is cumbersome but it can be handled. Several European research agencies are known to have used this procedure. A serious problem with this practice is that eliminating unlisted numbers threatens to bias results. In certain regions of some European countries, 30–40% of numbers – and especially new ones – are not listed. According to market research agencies females, singles, and widows are overrepresented among those having unlisted numbers.

Ideally, respondents should be informed about the study's purpose at the beginning of the data collection. Unfortunately, disguising the intention of a study is sometimes necessary for obtaining valid results. For instance, it does not make much sense to tell an employee in a supermarket at the first contact that the interviewer is a mystery shopper. This would render the research useless. In any case, the person whose behavior is under observation should be told about the study's true purpose immediately after the interview. And it should be possible for a respondent to have all available data (tape, video, written notes, etc.) withdrawn.

Interviewer – Agency Dilemmas

1. An agency is concerned about the performance of its telephone interviewers and conducts supervisory monitoring of their phone conversation with respondents.
2. A person applies for a job as interviewer at an agency and is hired for conducting in-home interviews. However, there is a hidden agenda: the person is employed by the secret service and by presenting himself as an interviewer of a respected and well-known agency, he obtains a unique opportunity of getting inside the apartment of persons being under surveillance for planning acts of terrorism. Notice that as long as the person does not reveal his true intention, the ethical problem concerns the interviewer. But if he reveals his true intentions while being hired, the problem is that of the agency.

Supervisory monitoring has been sharply criticized by Scandinavian organizations representing the interviewers' rights. A California law on eavesdropping, making it illegal to listen to an extension, may restrict the use of supervisory monitoring (Malhotra, 2007, p. 213). As long as the secret service agent acts in accordance with the country's law (presuming it is a democracy), and the pretended activity has been sanctioned by a judge, the moral question does not arise or it is offset or overruled by considerations regarding the security of the state.

Agency – Client Dilemmas

1. A client obtains multiple bids for a project. Since several bids appear to be of similar quality, he or she chooses the cheapest one. The study is carried out by the agency and the report is sent to the client. When the client asks for a slideshow presentation by the consultant and for some supplementary cross tabulation, he or she is told that this was not included in the original price and therefore a surcharge is required because of the extra cost. The problem is known as *low-ball pricing*. A decent agency consultant should provide a detailed break-down of the price-estimate of different processes/elements of the research study (say, questionnaire design x Euro, field interviewing y Euro, basic report z Euro etc.). A sheet containing such economic specification should then be made available to the client before the study is carried out. In this way the client gets a better overview of the added cost, say of new data analyses, slide show presentations etc.
2. In an effort to reduce costs, questionnaires from several clients are bundled and data for all clients are collected during the same interview. This can be a cheap way for clients to get primary data that otherwise would be too expensive. The method is known as an *omnibus interview* or *piggybacking*. An ethical issue arises in a case where the client believes that his or her study is carried out separately and thus is getting all attention from respondents. Had the client been aware that his or her, say, 20 questions were part of a 500-question interview, he or she might fear that the quality of the study would suffer because of the respondents' information overload. Sometimes, an agency works

for two close competitors at the same time without informing them about it. Since they are within the same business, there is a good chance that both clients are handled by the same consultant. Often, a successful study necessitates that the clients reveal confidential information like marketing strategies and competitive information. He or she would *definitely* not like any of this information to somehow be disseminated to the competitor. But what if the consultant discloses restricted data on one competitor to the other, say, with the purpose of doing the client a favor and thus strengthening the relationship? Handing over information to a competitor, given this setting, is unethical. However, assume that a potential client is told that the agency currently is working for a competitor, but that the two companies will be served by different consultants. If the client, nevertheless, wants to place his or her order, an ethical problem does not exist (another option would be that the agency has a policy of not working for two competitors at the same time). A special case is that of *syndicated research* studies like household panel surveys, retail audits, and TV-meter analyses (*see* PURCHASE, SCANNER, AND MEDIA PANELS). Here the client knows that he or she would have to share sensitive market information with a competitor, provided the competitor is willing to pay for this information. For instance, if the client is a coffee producer who tracks the markets shares of its brand portfolio, then there is no way to prevent a competitor from getting access to the same information.

3. An agency discovers that one of its interviewers has been faking most of his interviews. The interviewer confesses and is fired (after consulting its lawyer, management decides not to take the case into court). As a matter of fact, the person was by far the most productive interviewer of the agency. With regard to several recent studies for important clients he had conducted 20–30% of all interviews. One study was the last pre-election poll widely published showing that the government would be defeated. The government lost by a very close margin. On the basis of another analysis, management, after being much in doubt, decided not to launch the new product. Should the agency tell the clients that about one fourth of the data that were collected for the company are invalid? A decent manager would tell the clients what has happened, and maybe offer to conduct new interviews instead of using the faked ones, and make revised reruns for no charge. Also, clients could be offered a discount on new studies (however, the pre-election poll cannot be corrected!).
4. A food-processing company is about to launch a new product. So the marketing manager contacts the agency. His or her graduate marketing research textbook recommends a product development analysis called CONJOINT ANALYSIS. During the introductory interview, the consultant figures out that this technique – provided by the agency – is not appropriate given the client's problem. The brand is imported, is part of a standardized global campaign, and attributes like size, content, adverting, and price have been decided upon by the headquarters. When no attributes can be varied as part of the test, conjoint analysis is rarely the preferred option. Instead a target market segmentation by way of a CLUSTER ANALYSIS combined with an analysis of background data and media habits appears to be a better choice. Moreover, the former study costs about three times more than the latter. What should the consultant do? Should he or she go ahead with the conjoint study thereby getting closer to his or her own annual budget target (and give priority to the agency's revenues)? Or should the consultant recommend the better and cheaper solution? Normally, it is assumed that the agency should offer the best and cheapest solution to the client. Therefore, acting otherwise would be unethical.
5. A big management consulting company, T, selects a marketing research agency, Z, for carrying out a major customer satisfaction study for one of T'*s* clients. Some months later, the agency – having 200 employees – offers to send all its employees on a two-day training course at T once every second year – *assumed* that the consulting company would buy all its future marketing research

from agency Z. The deal is scheduled for four years and is then to be renegotiated. T agrees on the deal presuming that it automatically selects Z as supplier next time it needs to collect market research data either for itself or for one of its clients. Is the deal between companies T and Z unethical? The present scenario is rather complicated and disputed. Many researchers working within business-to-business marketing and relationship marketing do not at all regard such a deal as unethical. For these researchers, *reciprocity* is a useful business strategy that fosters "long-term relationships" and is "mutually beneficial." However, the US Supreme Court has issued critical comments. It called reciprocity in business relations *anticompetitive practice* and ruled that "The practice results in an irrelevant and alien factor intruding into the choice among competing products" . . . "It may increase the difficulty of winning customers for an improved product and . . . tends to discourage product and customer development work." For much more on the topic of reciprocity in marketing, see Tadajewski (2009).

6. A big retailer wants to establish a customer panel. So, he contacts a research agency and asks it to come up with an appropriate research design. The agency's consultant works full time for two weeks on all the specifics (screening, recruiting, weighting procedures, statistical algorithms, reporting standards, panel management, costs, etc.). He or she then sends a detailed prospectus to the retailer. For a while, he or she hears nothing. Some months later, he or she learns that the customer panel has been established. Moreover, he or she finds out that the experimental design follows his or her prospectus to the letter. There can be no doubt that the retailer either has established the panel on its own or has forwarded the prospectus to one of the agency's competitors and asked it to come up with a cheaper price, given the actual research design. This act by the client is clearly unethical and the case should be reported to the country's association of market research companies. Depending on the specifics, it may be illegal and the agency could go to the court. Some agencies today ask for a separate fee, say, 10–20% of the project price, for providing a specified bid. In that way, they are at least being compensated for the introductory work. The amount is then deducted from the project price if the company gets the order.
7. Presume that an agency is been contacted by a major bank, Q. The bank's marketing manager wants to set up an ambitious and comprehensive customer satisfaction index. The index should work as an early warning system for identifying customers who consider changing to a competing bank. The agency makes a research design that the bank accepts and the index system is established. A year later, when the bank manager meets with the manager of a competing bank, S, he or she is told that his or her bank is a client of the same agency and that agency has suggested that they should establish an identical satisfaction index. Since the agency probably has been charging an extra price for developing the index for bank Q, it is unethical to just reuse this for the competing bank, at least without informing Q. Notice that in this case, it is the *agency* that is behaving unethically.

Agency – Public Dilemma

About two decades ago, R.N. – an elderly marketing professor at a university of a big city in an EU country – wondered about the number of poll-based results published by an upstart opinion poll company. The company, D, was so far unknown and a short time after it was established, it began sending an astonishing number of press releases based on its polls, to the city's newspapers. The polls primarily dealt with actual political and societal issues. What surprised Prof. R.N. was this: from his many years of working at the university, he had learned that all polling companies he was aware of had a preference for using university students as interviewers. He regularly met his students at shopping malls, at street corners, and so on. Out of interest, he developed a routine of asking the student which polling company she/he was working for. At some point of time,

he began wondering why no one of his students was working for company D. He never met anyone. However, many of the published polls dealt with issues related to the city and the region. Since email was not available at the time, he could not send out a request to all the students. Instead he began asking hundreds of students on the campus but without success. Next, he called the company and asked them to provide him with some specifics about its interviewer base. But since interviewers had to be kept anonymous, they were not able to disclose any specifics.

Simple Minimum Standards

What follows is a set of minimum standards based on the marketing literature:

- The length of the interview should not be significantly understated. When contacting a person face-to-face, by phone or by email, one should not state that "it only takes fifteen minutes," when it actually takes about one hour. As a rule of thumb, an interview should not exceed 30 min. The exception is an in-home interview that can take up to one hour without overburdening the respondent (Malhotra, 2007, p. 325).
- Incentives for completing a questionnaire should be independent of the nature of the answer and should be kept within certain limits. If one projects the top acceptable value of a gift at $15 as suggested by the White House in 1964 to today, it corresponds to about €115 (Blankenship, 1964). In most cases, though, gifts will be in the range of €5–20. Note that a lottery among respondents is an exception ("A weekend trip for two to Berlin").
- Ideally, the respondent should be informed about the true purpose of the research as well as about the sponsor. However, in many cases it is necessary to disguise the objective to produce valid results (i.e., when doing mystery shopping, see above). If respondents are told at the beginning of the interview that the purpose of the study is to investigate consumers' attitudes toward fair-trade products and that the sponsor is the Max Havellar Foundation, then this may bias their responses. A way to handle this ethical dilemma is to inform respondents at the beginning that the purpose of the study and its sponsor are disguised. Additionally, one could tell respondents that the facts are revealed at the end of the interview. If the respondent cannot accept this, he or she can opt out. If he or she accepts, the interview is carried out, and immediately afterward the respondent is informed about the purpose and the sponsor (the procedure is named *debriefing*). If the respondent feels offended or misused his or her information is withdrawn (Malhotra, 2007, p. 244).
- If a respondent has refused to participate or has opted out, he or she should not receive any future unsolicited messages from the agency (mail, phone calls, emails, etc.).
- Basic information like sampling frame, sample size, contact method (phone, email, etc.), exact formulation of questions, computational algorithm, and so on, should be available either to the client or – if published in the press – to the public. Also, if a study is based on a mixture of methods, say, 40% in-home interviews and 60% phone interviews, then this should not be disguised. If a poll is 100% based on an Internet survey, it should not appear only as "a survey." Change in method across time should be made public. If a survey is conducted as a phone interview one year and as an email interview the following year, this methodological modification should be mentioned.

Marketing Research Code of Ethics

Owing to the wealth of ethical dilemmas in marketing research, it has been suggested that a formal certification for marketing research practitioners be established (Ferrell, Hartline, and McDaniel, 1998). To be effective, a certification should be issued by an authoritative institution like an industry association or a government. An example is the *Professional Research Certification* issued by the Marketing Research Association (the certification is given to an individual rather than to a company).

Academic researchers as well as agency managers have suggested setting up a code of decency for marketing researchers.

A code is a broadly written framework within which to determine when the behavior of a researcher is within or out of bounds. Codes can only cover principles of honesty while they cannot cover integrity. Integrity is more difficult to specify. Honesty is merely an avoidance of incorrect behavior, whereas integrity is a voluntary and positive form of honesty, where one takes initiative in acting decently. Basically, a code is a reminder of some principles those in the field are expected to follow (Blankenship, 1964). Unfortunately, violating ethical codes has no consequences for the offender. To make things worse, the greater the reward and the lesser the punishment for indecent behavior, the more likely it is that an individual will practice unethical behavior (Skinner, Dubinsky, and Ferrell, 1988).

Several organizations like the AMA (www.marketingpower.com), the Marketing Research Association (www.mra–net.org), and ESOMAR – the European Society for Opinion and Marketing Research (www.esomar.org) – have developed ethical codes or best-practice principles.

The codes, which can be downloaded from the organizations' web sites, address issues like the following:

- anonymity of the respondent and the right to opt out
- behavior of the interviewer
- description of research method (questionnaire, sample, etc.)
- specification of field work and of statistical analysis
- identification of the sponsor
- ownership of data
- confidentiality of results
- publication of findings
- responsibility for the research
- interviewing children and young people.

Being broad principles, the reader will find them to be of limited value as a practical guide or user's manual for handling new real-life situations. On the other hand, this is better than having no guidelines at all.

Several published studies investigate ethical issues in marketing research. Crawford, (1970) has been interviewing a sample of marketing executives and research directors. He found, for instance, that about 7 out of 10 questioned regarded the use of hidden tape recorders and one-way mirrors as unethical. However, disguising a study's true purpose was found acceptable by 80%. Akaah (1990) tries to replicate and improve the Crawford study. Although he uses a different method, his major findings appear to support the earlier study. Ferrell, Hartline, and McDaniel (1998) base their sample on AMA member firms. Like earlier studies, the researchers confront respondents with a list of hypothetical situations with ethical dilemmas. Their results indicate differences in perception of ethics between marketing research companies and corporate research departments. It seems that the awareness of ethical principles is somewhat higher in marketing research companies. Lund (2001) used the AMA directory as the sampling frame. He found that in case of a clash of interests marketing research practitioners regard pure ethical standards as being more important than what serves the company best.

While ethics plays an important role in marketing research, ethical principles and standards are also discussed within other marketing disciplines like marketing strategy (*see* ETHICAL MARKETING AND MARKETING STRATEGY) and international marketing (*see* GLOBAL MARKETING ETHICS).

Acknowledgment

The author would like to thank Niels J. Blunch for helpful comments on this contribution.

Endnotes

[1] In this essay we treat the term "*market or marketing research agency*" in a broad manner also including data subcontractors, university professors/students, and independent researchers. "Firm" and "company" are used as synonyms for agency. So, marketing research "entity" or "unit" might have been a better universal term, but to limit confusion we use "agency."

Bibliography

Akaah, I.P. (1990) Attitudes of marketing professionals toward ethics in marketing research: a cross-national comparison. *Journal of Business Ethics*, **9**, 45–53.

Blankenship, A.B. (1964) Some aspects of ethics in marketing research. *Journal of Marketing Research*, **1** (2), 26–31.

Crawford, C.M. (1970) Attitudes of marketing executives toward ethics in marketing research. *Journal of Marketing*, **34**, 46–52.

Dorroch, H. (1994) *Meinungsmacher Report: Wie Umfrageergebnisse Entstehen*, Steidl, Göttingen.

Ferrell, O.C., Hartline, M.D., and McDaniel, S.W. (1998) Code of ethics among corporate research departments, marketing research firms, and data subcontractors: an examination of a three-communities metaphor. *Journal of Business Ethics*, **17**, 503–516.

Lund, D.B. (2001) Deontological and teleological influences on marketing research ethics. *Journal of Applied Business Research*, **17** (2), 65–82.

Malhotra, N.K. (2007) *Marketing Research – An Applied Orientation*, Prentice Hall, Upper Saddle River.

Skinner, S.J., Dubinsky, A.J., and Ferrell, O.C. (1988) Organizational dimensions of marketing research ethics. *Journal of Business Research*, **16**, 209–223.

Sudman, S., Bradburn, N.M., and Wansink, B. (2004) *Asking Questions: A Definite Guide to Questionnaire Design – For Marketing Research, Political Polls and Social and Health Questionnaires*, Jossey Bass, San Francisco.

Tadajewski, M. (2009) The foundation of relationship marketing: Reciprocity and trade relations. *Marketing Theory*, **9** (1), 9–38.

exploratory research

Vikas Mittal

What is Exploratory Research?

As the name suggests, exploratory research in marketing refers to research that helps marketing researchers explore ideas, generate hypotheses, and formulate questions to be subsequently tested and verified in quantitative research. No wonder, exploratory research is typically conducted as a prelude to an informed, conclusive or descriptive research. Some examples of exploratory research include the following:

- A nursing home wanted to find out why its front-line staff had a high turnover rate. To understand what items to measure in the quantitative survey, several focus groups were conducted. They suggested that in addition to pay, respect from coworkers and supervisor, religiosity, self-efficacy, and work-family conflict were key factors affecting job satisfaction and turnover intentions. On the basis of this, items measuring these factors were included in the survey.
- A business school in the Midwest wants to conduct a survey to ascertain the demand for healthcare management programs. To design the various program profiles, a series of focus groups and in-depth interviews were undertaken with physicians, hospital administrators, residents, and nurses (*see* SERVICE INNOVATION MANAGEMENT; VOICE OF THE CUSTOMER).
- A health maintenance organization (HMO) conducted a survey of its members and found that 75% of its members want to be treated with respect by their physician. However, it was not clear what physician behaviors embody the concept of respect. A series of focus groups were undertaken to understand what respect meant to patients. They showed that the primary behaviors defining respect included physicians listening to the patient and not rushing the visit.
- A restaurant owner wanted to open a second branch in a nearby locality. Analysis showed that a weekly patronage rate of 300 guests would enable the second location to break even. Secondary research using census data showed that not only was the current population base too small to support such a volume, but also that it was likely to shrink in the future. On the basis of this, they decided not to conduct further survey research.
- A major pharmaceutical company wanted to improve the design of its key product – insulin kits. To understand how customers handled and used them, they videotaped different people using their product. On the basis of their observations of how people stored the kits, used them, disposed of them and where they used them, the design team gained new insights. They next held a focus group to further clarify and understand the reasoning behind the usage pattern. Insights from these were incorporated into the product redesign.
- A hospital wanted to understand peoples' reactions to its new website. Two variants of the new website (text heavy and graphics heavy) were launched. Users including patients, physicians, employees, and other constituents were encouraged to post comments online and also discuss them in an online chat room supervised by a focus-group moderator. By monitoring this online content, valuable insights about the website design were obtained.

In general, exploratory research consists of qualitative research and verbal, observational, and projective techniques. As described later, verbal techniques include focus groups and in-depth interviews, observational techniques include ethnographic research, and projective techniques include metaphor elicitation approaches. Sometimes, though not often, firms also analyze existing secondary data to explore possibilities and opportunities for further descriptive and confirmatory research. We focus here on direct-verbal techniques, focus groups and in-depth interviews, as they are the most heavily used forms of exploratory research in marketing.

How does exploratory research differ from quantitative research? In the marketing research industry, the terms *exploratory*

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

research and *qualitative research* are typically used interchangeably. In contrast, marketing researchers use the term *quantitative research* to refer to descriptive research as well as conclusive research. Descriptive research typically consists of a survey, cross-sectional or longitudinal, aimed at describing specific aspects of consumer perceptions, attitudes, and behaviors. Conclusive research usually consists of causal studies designed as an experiment to establish cause-and-effect relationships.

- Quantitative research relies on structured data collection obtained via surveys, secondary data sources (e.g., sales figures, media habits, and so on.) while qualitative research is designed to obtain unstructured data. In the former, for instance, rating scales may be used to ensure that data from all respondents is comparable, and subjected to statistical analysis. In contrast, qualitative research is, by design, unstructured whereby each participant does not respond to the exact same stimulus. This renders data noncomparable for classical statistical analysis.
- In quantitative research, structured data is collected using large sample sizes that are randomly obtained from a population of interest. Using the central limit theorem, then, findings from the random sample can be generalized to the population represented by the sample with a certain margin of error. For instance, in a survey of 200 patients who visited its emergency room in the past six months, a hospital finds 25% to be dissatisfied with the experience. The hospital may conclude with a certain level of confidence (say 90% or 95%) that among all the patients who visited the emergency room, $25 \pm X\%$ may be dissatisfied. In contrast, qualitative research relies on a small judgment sample that is carefully chosen to illuminate all aspects of the phenomenon of interest. Thus, participants in exploratory research should *not* be a random sample of the population of interest. Rather a *purposive* sample using the researcher's judgment should be used. The goal of selecting participants is to obtain maximum diversity in perspectives and opinions. For instance, the hospital may want to purposely interview patients who visited the emergency room for a variety of reasons ranging from serious trauma, mild fever, accidental poisoning, chest pain, and so on. Detailed in-depth interviewing may reveal differences in experiences enabling the hospital to get a better understanding of *why* people may be dissatisfied, and *how* the reasons for dissatisfaction may vary on the basis of the purpose of the visit.
- Stated differently, exploratory research helps to define the "what and why" while quantitative research helps to quantify "what percent." For instance, in the earlier example qualitative research allowed the business school to realize that, in addition to curriculum, the format of the offering was a critical decision factor. Further, physicians seem to favor a "weekend only" option over a "weekday-evening" option as the former was perceived to be more practical. The extent of preference, however, could not be ascertained. A quantitative survey of 200 physicians showed that 72% preferred the weekend option, while 28% preferred the weekday – evening option. This information enabled the school's dean to make informed choices.
- The above discussion should also clarify the commonly held misconception that qualitative and quantitative research techniques are rival alternatives to each other. "Should we do a focus group or a survey?" is a question that naïve marketing researchers often ask. This is unfortunate. Both research approaches complement each other and should be used to their fullest extent within the context of a comprehensive research program. As explained next, front-end qualitative research provides the building blocks for designing sound quantitative research, and back-end qualitative research provides insights that enrich the quantitative findings. To this end, seasoned marketing researchers, when presenting findings to the top management, will frequently combine both qualitative and quantitative research data during their presentation. Using both types of data not only allows one to draw statistically generalizable conclusions but also fully appreciate the "why" behind

them. In a presentation done for a HMO, researchers first presented the top two reasons for patient dissatisfaction with their visit to a primary-care physician, and these included "waiting time in office is too long" and "interaction with the physician." Anticipating questions that would arise during the management presentation, the marketing researcher conducted several in-depth interviews with dissatisfied patients to obtain insights into some issues

- Did you have to wait too long outside the office, or inside the examination room? Approximately, how long did you wait? Did anyone provide an explanation for the wait and/or apologize to you? What do you think is a reasonable waiting time?
- What do you mean by "interaction" with the physician? Did the physician listen to you attentively? Do you mean their demeanor and politeness? Were you able to fully understand what they were saying? What other aspects, for you, define physician interaction?

In this way, qualitative research should be viewed as an "insurance policy" that prevents the survey researcher from making important errors of omission and commission. An example of the first is a failure to include important attributes in a survey, while an example of the latter is a failure to use language that the respondent would easily understand, or use categories that are meaningful for the respondent. By incorporating qualitative findings, the researcher can make a much more enriching presentation, and stimulate discussion among the management than if he/she had just presented simple percentages from the survey results. In this example, the discussion about clarifying the meaning of "interaction with the physician" enabled the HMO management to make suggestions to conduct a follow up survey to gather additional information that could be used to address the issues identified in the original survey.

Comparative Overview of Exploratory Research Techniques

At the broadest level, exploratory research techniques can be classified into three categories: (i) verbal-direct, (ii) nonverbal/indirect, and (iii) observational techniques. These are summarized in Table 1.

Verbal-direct techniques are those where the researcher directly asks questions of the participants who in turn verbalize their responses to provide insights. These include focus groups and in-depth interviews. These techniques assume that the participant has the ability and motivation to verbalize their feelings, judgments, thoughts, and intentions. In many cases, such as when discussing concepts that are radical or completely new, or when people do not fully understand their own behavior, these assumptions may be violated. Relative to observational techniques that require elaborate ways to observe and record usage behavior in naturalistic settings, focus groups and in-depth interviews can be conducted with relative ease and are much cheaper and faster. Perhaps this is why they are also the most popular.

Nonverbal and indirect techniques are best suited to surface those issues that respondents are either unable or unwilling to verbalize upon direct questioning (*see* PROJECTIVE TECHNIQUES). In one popular technique involving metaphor elicitations (Zaltman and Coulter, 1995), respondents are asked to collect various pictures from magazines that describe their feelings of a particular phenomenon. Then, a trained interviewer can help them to connect the bridge between deep-unrealized thoughts and verbal articulation of the same. For instance, when asked about the recent stock market turmoil, some investors were unable to articulate how they felt. They were then asked to collect pictures that best represent their feelings. One respondent brought a picture of a boat capsizing in a storm. Under the guidance of a moderator, the respondent then was able to articulate his/her thoughts that he/she felt (i) rudderless, without direction on how to navigate the financial turmoil, (ii) abandoned, and (iii) she wanted more stability.

Ethnography or observational research relies on the direct, unobtrusive observation of people in a naturalistic setting using the product or service of interest (*see* ETHNOGRAPIC RESEARCH; PERSONAL OBSERVATION; OBSERVATION METHODS). It is typically most useful in designing products such as electronics

(e.g., handsets, keyboards, TV, and computer monitors) and other products (e.g., cars) where the usage experience critically depends on sustained interaction between the product and the person (Anderson and Narus, 1998). The resulting experience is captured in the form of pictures, videos, field notes, and drawings and then used to draw conclusions. Observational research is also used in retail settings where consumer patterns in terms of walking the aisles, browsing, and shopping may be captured using a video and then analyzed.

Key Steps in Conducting a Focus Group or IDI

Both focus groups and IDIs are the most widely used form of exploratory research (*see* FOCUS GROUPS AND DEPTH INTERVIEWS). In conducting a focus group (*see* FOCUS GROUPS AND DEPTH INTERVIEWS), marketing researchers typically follow a series of steps, each step representing a key decision. A brief outline is provided here.

Who should participate? It is useful to have participants who have direct involvement with the issues that are being investigated. These typically include the end users of the product or service under consideration. Other participant groups typically used in marketing research include current customers, prospective customers, lost customers, front-line employees, and company executives. Sometimes, to fully understand competing perspectives, it is also useful to involve those whose involvement is indirect, but who can provide a different take. Thus, in understanding employee satisfaction with their job, a firm also interviewed immediate supervisors to understand the phenomenon from the supervisors' perspective. In other cases, important decision makers (e.g., in the purchase decision) are considered, even though they may not directly use the product. In designing toys targeted toward 9- to 12-year-old kids, the toy company not only held focus groups among children, but also a focus group among parents, the latter being key decision makers. The involvement of multiple stakeholders and their

Table 1 Qualitative research: a comparative view.

Focus Groups/IDI's (Verbal–Direct)	*Projective/metaphor (Nonverbal, Indirect)*	*Ethnography (Observational)*
Ask people to talk about their behaviors and motivations	Typically people are given a set of pictures and asked to make a collage from them	Observe people's behaviors and motivations
Artificial setting and moderator driven	Semi-supervised. Done at facility or at home	Naturalistic setting embedded in context
Semistructured and follows a script given to the moderator	Open-ended and varied	Open-ended and varied
Respondents articulate their own feelings in their "own words"	Respondents produce a collage from the pictures that best highlight their understanding of the issue/construct	Observers impose their vocabulary and interpretation on respondent's behavior
Cheaper to do	Cheaper to do	More expensive to do
Data: transcript of discussion	*Data*: collages by respondents	*Data*: photos, videos, drawings, field notes, tape-recordings, and so on
Focus groups, IDI	*Metaphor elicitation*	*Field observation*

opinions becomes more and more important in business-to-business marketing situations and in multistakeholder industries like healthcare.

What should be discussed? Though unstructured, focus groups and IDIs are not free flowing conversations. Rather, the discussion follows a "moderator's guide" or an "interview guide." This is an outline of the topics that need to be covered, and also gives an approximate idea of the time to be spent on each topic. The guide is usually finalized by the moderator, in consultation with the client. Importantly, the moderator or the interviewer may not follow the outline exactly. Rather, an experienced moderator will likely jump back and forth between topics, depending on the flow of the actual focus group. Further, for selected topics the participants may be probed to clarify and provide examples. In other words, the guide provides a roadmap for the discussion.

It should be noted, that the discussion guide for later focus groups or IDIs can differ significantly from earlier ones. Thus, as the research team feels it has gained sufficient insight into a particular topic, the topic may be removed from subsequent focus groups and be replaced by an issue about which further clarification is needed. Sometimes, for example, an issue may surface toward the end of a focus group, and enough time is not available to discuss it. In such cases, the moderator may bring up the issue in a subsequent group to gain insights.

How long should it last? A typical focus group lasts from 60 to 90 minutes, and a typical in-depth interview lasts from 45 to 90 minutes. During this time, the moderator will greet and introduce the participants, orienting them to the "rules" (e.g., not talking over each other and so forth). After that, the discussion proceeds as described in the previous section. While in some instances the interaction can last longer, it is advisable to limit them to no more than 90 minutes. Typically, by that time participants can feel fatigued and may lose the motivation or interest to fully participate in the discussion.

How should the participants be incentivized? Usually, marketing research firms provide both cash and/or noncash incentives, and the magnitude of the incentive can vary depending on the participants. In one focus group of car dealerships, the research firm donated $500 to their favorite charity to induce participation. In another focus group among parents of a local Montessori school, a group of MBAs provided food and beverages to thank parents for their participation. In another focus group of nursing aides, an incentive of $50 was used to recruit participants. An online focus group designed to understand perceptions of different websites promised a chance to win a $100 gift card for an online shopping site. This illustrates the wide range of incentives that can be offered to secure participation.

How many participants? A typical focus group has between 5 and 10 participants, with an average of seven participants. Research has also shown that about 7–9 focus groups and 20–25 in-depth interviews are enough to surface over 85% of all the customer needs. As explained earlier, researchers should ensure that focus groups and IDIs are conducted in a serial fashion, with later attempts building on and clarifying information obtained in earlier interactions. For a typical focus group, 10–12 participants may be recruited. Accounting for no-shows and cancellations, a desired size of 7–10 participants results.

It is also typical to collect some demographic data on the participants to get an idea of their background. Thus, upon arrival at a focus-group facility, potential participants may be given a short survey to determine eligibility and record basic information. Usually, this is also accompanied by appropriate disclosures to the participants. For instance, if the focus group is being taped or the conversation recorded, the participants should be clearly made aware of it, and preferably their consent should be obtained in writing.

How should the results be described? A key output is the actual transcript of the entire conversation. It is customary to include a video or an audio tape of the focus groups or IDIs in an appendix. Some companies prefer to have a written transcript of the conversation. However, these can run into several hours or hundreds of pages, and are not easily analyzable. Thus, the market researcher will typically use the transcripts to develop key themes that describe the main conclusions of the

focus groups/IDIs along with illustrative quotes. For instance, the theme of "lack of respect" was supported by quotes such as the following, in a focus group of nursing aides working in elder-care facilities: "Doctors were very rude to the aides. They treated everybody horribly," "it's like you're at the bottom of the pole" and "I'd like it if there were more respect for aides". Doctors and nurses should treat us like partners and rely upon our knowledge of the patients." There are software packages designed to conduct textual analysis and surface themes, though a discussion of these is beyond the scope of this article (*see* CONTENT ANALYSIS). Nonetheless, organizing the findings of qualitative research around key themes with quotes illustrative of the themes is the most common way of describing and presenting the findings.

In some cases, researchers may be tempted to describe the findings quantitatively. For instance, on the basis of a focus group of six patients, one naïve researcher tried to draw the following conclusion: because "quality of care" was mentioned 23 times during the focus group, it is more important than "doctor makes the correct diagnosis" which was only mentioned three times. The number of times something is mentioned in a focus group is a function of many things, and none of them may be related to its importance. Consider the following possibilities:

- There may be one participant in the focus group who was very verbose and dominated the conversation. This participant may have discussed quality of care at length, perhaps because he/she had a particularly good or bad experience with that aspect of care.
- It may be that all the participants agreed that "doctor makes correct diagnosis" was the most important attribute, but they disagreed about the meaning of "quality of care." Thus, a lot of time was spent trying to understand what "quality of care" meant to different participants.
- It may be that the physician group commissioning the focus group only wanted to learn about "quality of care" and put that on the moderator's guide, and that was why it was mentioned many times.

If the researcher is truly interested in understanding the relative importance of these two attributes, then this focus group should be followed up with a quantitative survey. In the survey, respondents could rate all attributes on a 7-point scale, (7 = very important, 1 = not at all important). From that survey, suppose 75% give "doctor makes the correct diagnosis" a rating of 7, and 62% give "quality of care" a rating of 7. Then, using a z-test of proportions a hypothesis test could be conducted, and conclusions about their relative importance to the patients can be drawn.

Should we pay extra for an experienced moderator? Within the constraints of a budget, the answer to the above question is yes, yes, and yes. The importance of a high-quality moderator/interviewer is frequently compared to the importance of a high-quality lawyer at a jury trial. Though it seems easy on the surface, moderating a focus group requires interpersonal, social, organizational, and verbal skills that may take years to develop. For instance, rather than being overly talkative and/or effusive, good moderators fade into the background while directing the focus-group participants and the flow of the conversation. They are firm, particularly with the occasional problem participant, but in a way that does not affect others. Similarly, they are adept at ensuring that the shy or quiet participant's views are brought forward and incorporated in the discussion. They also take time to fully understand the client's needs, translating them into a topic-guide and then ensuring that all aspects of the guide are covered in the discussion, as well as in the final report. Not surprisingly, high-quality moderators are in high demand frequently having scheduling backlogs running into weeks if not months.

Should we conduct focus groups or in-depth interviews? Both focus groups and IDIs are direct-verbal techniques designed to elicit the judgments and opinions of the participant (*see* FOCUS GROUPS AND DEPTH INTERVIEWS). When studying sensitive topics such as drug use, cheating, personal hygiene products, and cosmetic surgery, some researchers agonize whether they should use focus groups or

in-depth interviews. On the one hand, some make the argument that an in-depth interview affords respondents privacy because the respondent is alone and can therefore confide in the interviewer. Because of this the respondent can be more forthcoming. In a focus group, they argue, social desirability may distort the respondent's responses. On the other hand, it has been argued that a focus group allows people to feel at ease and put forth their ideas when they see others being forthcoming. Over the years, people have designed many studies to directly compare the outcome of conducting focus groups versus a series of in-depth interviews. Findings show that both techniques yield similar information, with neither technique holding an advantage over the other. What are the reasons for this?

A well-trained focus-group moderator or a well-trained in-depth interviewer can both put respondents at ease and guide them to explore ideas, irrespective of the sensitivity of the topic being discussed. Thus, privacy afforded by an in-depth interview or synergy afforded by a focus group is a poor argument for choosing either. Rather, in practice the choice is driven by the logistics of the project (e.g., timeline and resources) and respondent convenience. For instance, finding that it was cost prohibitive and virtually impossible to invite a company's CEO to a central location, the marketing researcher decided to conduct in-depth interviews via telephone to understand sustainability initiatives at these companies. In many cases, a mix of focus groups and in-depth interviews is used. For instance, a series of in-depth interviews can be done following a focus group to gain further insights into the issues that surface in a focus group. By staggering the interviews, ideas obtained in previous encounters can be elaborated upon in subsequent ones. In another study, researchers wanted to understand factors that can improve or impede retention in high-stress jobs such as taking care of the elderly in nursing homes. They wanted to reach out to CNAs (certified nurse aides) both in urban and rural areas. Because of the proximity of several nursing homes in urban areas, it was easy to conduct focus groups. However, owing to the inability of CNAs in rural areas to travel long distances, telephonic interviews were conducted. Later, a comparison showed remarkably similar results for participants in urban and rural locations, and thus, data from both methodologies were combined for the sake of analysis and insights.

Complementary Use of Exploratory and Quantitative Research

Sometimes qualitative research is done as stand-alone research to gain an in-depth understanding of a phenomenon, such as customer reactions to a product concept or advertising copy. Frequently, however, it is used in conjunction with quantitative research, most notably survey research. In this manner, the use of qualitative research can be subdivided into three parts, each corresponding to different stages of the survey research

- *Front-end exploratory research*: One main use of exploratory research is to generate survey content such as key categories of constructs to be included, survey items to measure the constructs, and specific language corresponding to each construct. For instance, a bank wanted to design a customer satisfaction survey to measure satisfaction with customers' visits to the bank branches. Observation of several branches and focus groups with customers identified several components, including waiting time, interaction with tellers, products and services, and fees. Then, specific items corresponding to each area were developed by studying the focus-group transcripts. Qualitative research can be very useful to understand the vocabulary customers use to describe their experiences and/or the firm's brand. For instance, a company interested in pricing research wanted to design a survey to measure the likelihood of purchase at different shampoo prices per ounce. They did not know how to frame the pricing question for the survey. Qualitative research showed that the customers thought in terms of the price of a bottle of shampoo (small, medium, and large) and not per ounce. On the basis of this insight, the survey was designed to incorporate the customers' mindset, and enabled the firm to gain valuable insights.

- *In-process qualitative research*: Frequently referred to as "*pretesting*" and "*cognitive testing*" of a survey, in-process qualitative research is used to refine an initial draft of a survey (*see* PRETESTING A QUESTIONNAIRE; QUESTIONNAIRE DESIGN). The survey is administered to a respondent with a researcher monitoring the respondents' verbal and nonverbal behavior. For example, if the respondent takes too long to answer a question, it may be that the question is ambiguous and/or the response categories do not make sense. Sometimes the respondent may leave a question unanswered. The researcher can then ascertain the reason for nonresponse which can range from unclear question wording, to multimeaning response options, to offensively framed questions. The researcher can then probe and clarify such issues, to revise and refine the survey accordingly.
- *Back-end qualitative research*: Once a survey is completed, sometimes a focus-group or in-depth interviews can help clarify observed patterns of results, or the logic behind them. For instance, a physician found that 35% of his/her patients viewed his/her as "nonresponsive to my needs." A series of in-depth interviews showed that patients felt he/she (i) did not acknowledge their greeting in his/her office; (ii) was seen as unwilling to give referrals, and (iii) frequently made patients wait over 30 minutes in the office. The physician decided to address these issues to improve his/her perceptions of being responsive to patient needs.

Misusing Qualitative Research

There is also a potential danger that qualitative research may be misused and abused. There are many occasions and instances where qualitative research should not be used. In particular, misuse of qualitative research can occur if qualitative research is used as a substitute for quantitative research. Often, people will cite qualitative research to make assertions such as "attribute '*X*' was mentioned by many respondents in the focus group. Therefore, it must be very important for our customers." As explained before, the only conclusion one should be drawing is that "attribute '*X*' that was mentioned in the focus group should be part of the quantitative survey." An attribute may generate a lot of discussion for many reasons and not necessarily because it is important. In some cases, managers' misuse qualitative research by selectively using quotes from respondents to push their own agenda/viewpoints. This should be avoided at all costs by sharing the actual transcript (typed or as a video file) with everyone and fostering a detailed discussion of the findings.

Finally, one should resist the temptation of "quickly throwing together" a focus group to get "some idea" about consumer needs. On the contrary, just like quantitative research, qualitative research should be carefully planned in terms of the content of discussion (e.g., moderator's guide), participant selection, venue, transcription of data, interpretation of data, and the nature and strength of conclusions drawn. By following proper guidelines one can not only maximize the impact of qualitative research, but also ensure it is not abused by others.

Emerging Trends

The spread of the Internet has also transformed qualitative research. As boundaries of time, distance, and content are blurred exciting new avenues for capturing data for exploratory research have opened (*see* WEB SURVEYS). For example,

- Online focus groups allow participants to exchange ideas under the guidance of a trained moderator. Participants log into a secure chat room where the online moderator can conduct the focus group. As video conferencing technology proliferates, distance-based focus groups will be further transformed.
- Companies are paying attention to online blogs (from its customers and others), online comments (e.g., epinions.com), and other sources which provide useful exploratory research. Insights from these sources can guide the design and focus of quantitative research to a large extent.
- Online videos and posts from interested customers can be rich sources of observational research. Ethnography will become

more prevalent as customers become more open to allowing companies to remotely observe product and brand usage in real time.

- Intranets within a company have become popular for eliciting employee comments to gain qualitative insights and feedback. For instance, initial product testing can be conducted with the employees. Sometimes, they may be asked to provide feedback on branding initiatives such as the redesign of the company's website.
- Finally, researchers have also begun to quantitatively analyze large sources of qualitative data. For instance, Yahoo users post thousands of movie reviews online. These can be quantitatively analyzed to understand various aspects of movies that can predict viewer tastes.

Concluding Comments

Exploratory research is the foundation upon which the success of a marketing research effort is built. It should be an integral part of any marketing research effort, and not an afterthought. We hope that this article will provide the reader with a basic foundation for conducting exploratory research to gain marketing insights that may drive organizational success.

Bibliography

Anderson, J.C. and Narus, J.A. (1998) Business marketing: understanding what customers value. *Harvard Business Review*, 53–67.

Mittal, V., Rosen, J.M., and Leana, C. (2009) A dual-driver model of turnover and retention in the direct care workforce. *The Gerontologist*, forthcoming.

Zaltman, G. and Coulter, R.H. (1995) Seeing the voice of the customer: metaphor-based advertising research. *Journal of Advertising Research*, **35** (4), 35–51.

marketing research process

Gaurav Bhalla

Introduction

Capital One is one of the largest issuers of credit cards in the United States. However, Capital One is more than just a credit card company. Capital One is also a retail branch banking company. A few years ago it acquired Hibernia, North Fork Bank, and Superior Savings of New England, N.A., which gave it a footprint of over 700 branches and 1100 ATMs in New York, New Jersey, Connecticut, Virginia, Texas, and Louisiana.

Let us consider 2004, before the acquisitions took place and imagine that the Executive Team is discussing the diversification from credit cards to retail banking. The CEO wants to know if Capital One's brand equity from credit cards will extend to banking. Most of the executives in the room are positive that it will. However, the CEO requests that a formal marketing research (MR) study be commissioned to get the customer's perspective.

As you can imagine, to comply with this request, the MR professionals at Capital One will have to make a series of systematic and related decisions, such as the following:

- Agreeing on the specific marketing question to be addressed.
- Translating the marketing question into specific research objectives.
- Deciding on what type of data to collect and how to collect it.
- Identifying the subpopulations to be included in the sample: all potential retail banking customers or just credit card holders, and so on.

As the mini case study illustrates, one of the defining characteristics of contemporary organizations is the use of formal information to make sound decisions. Gut feel and intuition are necessary, but not sufficient. The CEO could have gone with a majority vote, but decided instead to supplement his/her team's judgment with the market data. MR is one of the most frequently used methods to formally collect market data. For an MR project to yield the information desired, it is important that the professionals in charge of executing the project follow a systematic method, as discussed in the mini case study.

The traditional MR literature refers to the sequential systematic steps taken by MR professionals for the execution of MR projects as the marketing research process (MRP). The rest of this article will discuss this traditional MRP and its various constituent steps in detail. Following that, the article will discuss certain modifications and adaptations to the traditional model, on the basis of how the MRP may actually play out in practice. Lastly, the article will also discuss on how the emphasis and focus at a number of companies is shifting away from the transactional aspects of the MR process to thinking of it as a business process, where the goal of MR is not merely to provide results, findings, and insights, but to influence business outcomes.

Overview

Most common and popular MR reference sources describe the MR process similarly (Kumar, Aaker, and Day, 2009; Malhotra, 2007; Zikmund and Babin, 2007). Broadly speaking the MR process comprises of three main phases (*see also* MARKETING RESEARCH PROPOSAL):

- Predesign or planning phase (*see also* CRITERIA FOR EVALUATING SECONDARY DATA)
- Design phase (*see also* EXPLORATORY RESEARCH; EXPERIMENTAL DESIGN; FOCUS GROUPS AND DEPTH INTERVIEWS; OBSERVATION METHODS; SURVEY RESEARCH)
- Execution or implementation phase (*see also* WEB SURVEYS; FOCUS GROUPS AND DEPTH INTERVIEWS).

The focus of the predesign or planning phase is to ensure that all stake holders are in agreement on why the research is being conducted, which business questions will be addressed by the research, the specific research objectives that will guide the data collection, and whether or not the data proposed to be collected will incrementally influence business decision-making. The intent is to establish clarity of purpose, develop

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

focus for data collection, and ensure that the information collected will deliver incremental value to the decision maker.

The design phase is concerned with issues related to the best, or most suited, approach for data collection, and with tactical details concerning data collection, such as questionnaire design, types of questions that should be asked, and the types of analysis that would help extract the maximum information from the data. Decisions concerning the best, or most suitable, approach for data collection are always taken in the context of the purpose and focus established in the predesign phase. Tactical decisions are constrained by the overall research approach preferred; they also influence decisions concerning the overall approach. Take the example of a concept test. If data are being collected over the telephone, it would be difficult to administer questions that require real time evaluation of visual stimuli. If the client believes that without evaluating pictures of concepts obtaining accurate respondent perceptions will be difficult, then the telephone data collection approach will have to be modified.

The execution or implementation phase deals entirely with the collection, analysis, reporting of data, findings, conclusions, and insights. This phase also involves making recommendations to the decision maker, on the basis of the insights acquired, concerning alternatives for action and their relative merits.

Properly conceived and executed, all three phases collectively determine the overall quality and value of the research conducted. It is not useful to discuss which phase is more important, owing to the interdependence that exists between the three phases. To illustrate, if the business and/or research problem is inaccurately or incompletely stated, no matter how brilliant the research design, the outcome of the research will not be helpful to the decision maker. Similarly, errors in designing the research can reduce the overall quality of the research, no matter how brilliant the execution. And finally, the best planned and designed research can be defeated by flaws in execution – poor sample quality, ill-trained interviewers, and weak analysis can all detract from the overall quality of the research.

Figure 1 depicts the three macrophases of the research process and the key decisions within each phase.

Predesign or Planning Phase

Management thinking is not unlike fashion. Certain ideas catch on and then dominate the consciousness and agendas of companies, before giving way to the next. Several candidates come to mind when one looks in the rear view mirror – total quality management (TQM), customer satisfaction, six sigma, customer loyalty, customer relationship marketing (CRM), and so on. They have all had their time in the sun and then faded into the background.

Customer focus is today's differentiator. Every company today claims to be customer focused; yet when you experience the company, its products, its people, and its service, the impression that you get is quite contrary. Imagine you are the chief marketing officer (CMO) of an organization that operates at the intersection of customer service and retail – like Home Depot (world's largest home improvement retailer) or Best Buy (world's largest specialty consumer electronics retailer). You are well aware that your company claims to be very customer focused, to provide unimpeachable service, and to always put the customers' interests over the company's interests. Yet, you have heard from friends and family and from newspaper reports that most consumers think that companies are not customer friendly and that their customer service stinks. You are also aware that if you google for the term 'customer service stinks' you will get thousands of links, blogs, posts, news articles, books, and so on, complaining about companies and the service they provide.

Clearly, there is a disconnect between how the company views its service offering and what the market believes it is getting. Can MR help the CMO? Where should the CMO begin – internally with front-line employees, or externally with disgruntled customers? But is being disgruntled what it is all about? The CMO has also heard that not everyone who receives poor service complains. But, since the company is committed to providing excellent service, technically speaking, every customer who did

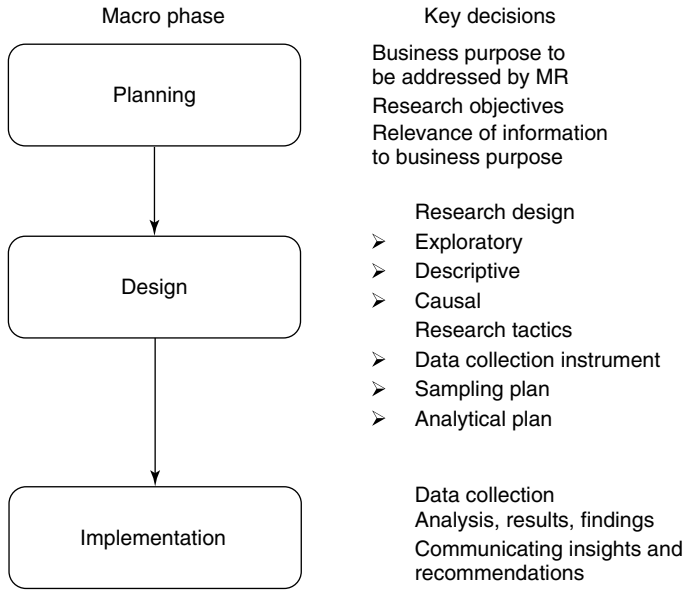


Figure 1 The marketing research process.

not perceive receiving excellent service reflects underachievement on the part of the service department. How should the CMO proceed to identify the nature and sources of this underachievement and then, to root it out?

Hopefully, the mini case study above should demonstrate that decision making in the business world is not always straightforward or easy. Business situations are messy and multidimensional, making it difficult to decide the subset of issues to focus on and their prioritization. It is very easy to imagine two MR professionals interacting with the CEO in the above scenario, one recommending that the company spend time gathering more facts before conducting formal MR and the other recommending proceeding with formal MR as all the facts that are needed are already available.

To be certain that significant aspects of the problem are being addressed and that following the research decisions can be made without the need to collect additional data, the following steps should prove helpful.

- No matter how complex the problem/opportunity, it always helps to generate a set of hypotheses or speculations as to the reasons contributing to the problem/opportunity. If there is uncertainty concerning potential contributors, making it difficult to hypothesize, then appropriate investments should be made in fact finding and desk research to learn more about the relevant dimensions of the target problem/opportunity.
- Assuming enough is known about the service problem discussed in the case study, a few potential hypotheses could be the following:
  - Poorly trained service representatives.
  - Ambiguous processes on how to handle customer complaints.
  - Inadequate communication with the customers on service options and complaining procedures.
  - Unduly long wait, hold, and lead times in handling customer complaints/returns.
- No company has enough time or resources to research all issues that it is uncertain about. Dealing with uncertainty and risk of incomplete information is par for the course in today's organizations. Once the set of hypotheses has been generated, managerial judgment and commercial astuteness will need to be exercised in making trade-offs

and determining a subset of hypotheses to be researched.

- A forward–backward approach can help in making these trade-offs. For example, the CMO can telescope forward to the point of decision making and articulate the information inputs that will enable him/her to make an effective decision.
- The MR professional can now work backward from these information inputs to the hypothesized reasons to identify the subset of hypotheses that should determine the focus of the research. As indicated above, if the subset of hypotheses to be researched is unduly large, trade-offs will have to be made.

The iterative thinking recommended above has one purpose – to make certain that the business problem being addressed is relevant, given the situation the company finds itself in. The ability to get to the heart of a problem is an invaluable skill. This skill can be nurtured and developed by investing in heuristics that emphasize iterative thinking, as recommended above, till it becomes second nature.

In the predesign phase, most standard textbooks of MR discuss the need to estimate the value of information, as all MR has a cost attached to it. The idea is that if the decision can proceed with certainty, independent of the research information collected, then the proposed information has zero value. This author has no desire to discount concepts like the expected value of information. However, in the day-to-day press of the commercial world, the determination of the incremental value of information is essentially subjective. To a very large extent, the decision maker and his/her key cohorts determine whether the information is likely to be valuable or not.

Design Phase

The key objective during this phase is to develop a blueprint for data collection, with the overall goal being the fulfillment of the study's business and research objectives. The decisions made during this phase are both conscious and iterative. Decisions made earlier in the research process, such as objectives and scope, influence subsequent decisions in the MRP such as approach and analysis. However, decisions concerning data collection and analysis can also influence decisions made previously. For example, lack of broadband Internet access among certain sections of the population may lead to a revision in research design; purity of design, emphasizing a single method of data collection may be sacrificed to accommodate a larger cross section of the population.

Here, just as in the case of the overall MRP, the MR literature is unanimous in its classification and discussion of major research approaches. Most texts classify research into three categories – exploratory, descriptive, and causal. Descriptive and causal can be further subclassified; for example, descriptive research can be either cross-sectional (one sample at one point in time), or longitudinal (repeated samples at successive points in time). Figure 2 summarizes the main and subcategories of research design.

For the sake of completeness, it is important to point out that some MR text and reference sources use a different classification system. Brief details are provided below.

- Research approaches are broadly classified in two categories – qualitative and quantitative.
- Qualitative research approaches are then subclassified into categories, such as in-depth interviews, small group research (focus groups), observational research, and so on.
- Similarly, quantitative research approaches are subclassified into categories such as descriptive, evaluative, predictive, and so on. The subcategories may themselves be stratified further. For example, there are instances where predictive research is classified further into experimental design and modeling research.

Additionally, there are specialist designs that are not discussed in this article or in most MR text books (*see also* ETHNOGRAPIC RESEARCH; MOTIVATION RESEARCH; OBSERVATION METHODS; PERSONAL OBSERVATION; PROJECTIVE TECHNIQUES). For example, qualitative researchers may prefer the narrative research approach, which analyzes respondents' first person stories recorded as narrative text

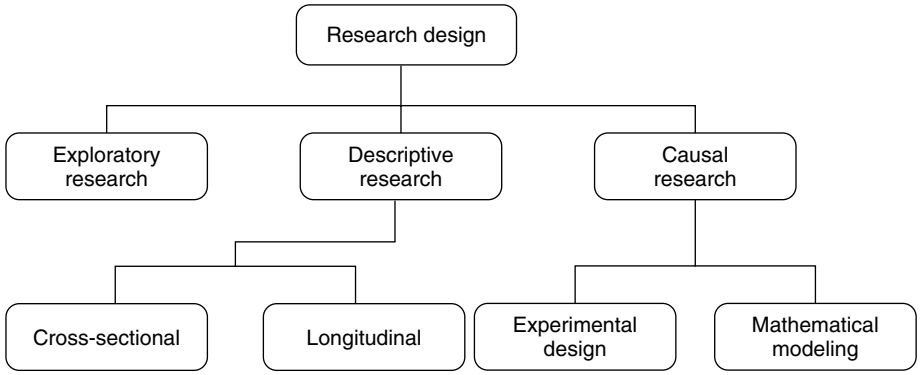


Figure 2 Taxonomy of research designs.

(Clandinin and Connelly, 1989). Similarly, given the increasing popularity of social media, quantitative researchers may prefer to use specialized agent-based modeling techniques to study the spread of influence and the diffusion of innovations (Bonabeau, 2002). Generally speaking, specialized qualitative research approaches are usually borrowed from psychology, sociology, and anthropology, while specialized quantitative approaches are usually borrowed from disciplines like econometrics, statistics, and computer science.

Lastly, even though this section and MR text books discuss the various approaches to data collection as separate activities, power in research design is often derived by combining research approaches. For example, researchers interested in modeling the breakfast food choices of different ethnic groups (Hispanic, Asian Indian, African American, White Caucasians, and so on.) may use qualitative research to first understand the context variables driving choice (family size, budget, time, perceived importance of breakfast, food preferences – cold, hot, freshly cooked, precooked, and so on). This could then be followed up by a data-collection-modeling phase, which would gather descriptive data from a larger number of respondents and then model it using statistical methods like Logit modeling (*see also* LOGIT MODEL).

Execution and Implementation Phase

The execution and implementation phase is concerned with data collection, analysis, and communication of results, insights, and recommendations.

Traditionally, in-person, mail, and telephone methods have dominated MR data collection. With the growing popularity of computers in the 1980s, computer aided interviewing grew in importance for both the in-person and telephone formats. Since the mid-1990s, with the increase in ownership of PCs and in Internet usage, the share of web-based methods of data collection has increased significantly. Figure 3 presents a classification of key data collection options available to a researcher.

Activities related to data collection, including questionnaire writing and analysis have been significantly transformed by the availability of specialized software. The goal of data collection software, like Confirmit, MarketTools, and Surveywriter is to help make the data collection phase efficient and painless, by passing and assigning the drudgery and routine elements to the software.

The goal of the analysis software (SAS, SPSS, and so on), on the other hand, is to disseminate statistical computing power as widely as possible; computing power was once available only to those with specialized training. As a result, lack of knowledge of how the statistical procedures work is no longer a barrier to designing specific types of research. For example, an analyst does not need to know the mathematical and/or statistical intricacies of utility estimation to design a conjoint study. Admittedly, more complex studies requiring complicated experimental designs will still require relevant

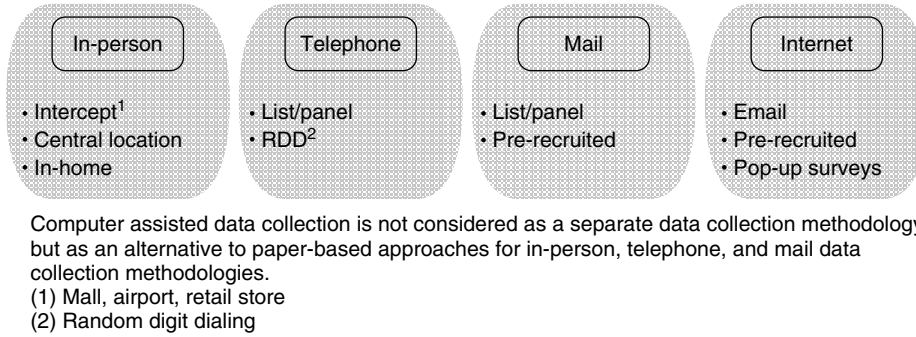


Figure 3 Classification of data collection methods.

knowledge, limiting the number of research companies able to offer the service.

Fortunately, knowledge still plays a role in designing research, in analyzing data, and in drawing meaningful insights from the analysis. However, at the cost of oversimplifying, it can be safely said that the widespread availability of relevant literature, MR courses/training, software, and other technological aids have made the competitive space for MR services much flatter than it once was.

Statistical procedures, both univariate and multivariate, still form the backbone of most of the analysis conducted on MR data (*see also* CLUSTER ANALYSIS; CONJOINT ANALYSIS; EXPLORATORY FACTOR ANALYSIS; MULTIDIMENSIONAL SCALING OF PREFERENCE DATA; DISCRIMINANT ANALYSIS FOR MARKETING RESEARCH APPLICATIONS; ANALYSIS OF VARIANCE AND COVARIANCE; MULTIPLE REGRESSION). Over the years, as the demand for predictive research has increased, this repertoire has been supplemented with econometric and statistical procedures that emphasize nonlinear estimation and analysis procedures. Additionally, the new procedures also allow for simulation of results to allow the sponsor of the research to test and evaluate a variety of decision scenarios, without the need to collect new data.

Following the completion of an MR project, it is customary for the agency or the department that has conducted the research to share:

- results and insights with the sponsors and key stakeholders who commissioned the research and;
- make recommendations about possible courses of action, on the basis of the results of the study.

PowerPoint presentations are the mainstay here. Few MR presentations use a truly multimedia platform to communicate results; in a few instances, where qualitative or ethnographic research is conducted, videos of respondents in their natural environments sharing details relevant to the study may be shared with the sponsors of the study. In some instances, following the presentation of results, work groups may also be set up to determine what actions should be taken on the basis of the insights acquired, and what future research needs to be done, if any, to support future decision making.

Marketing Research Process in Practice

The traditional approach, as presented and discussed in this article and other MR texts, reflects a linear, sequential, rational approach to decision making. Like all models, it captures only a portion of reality. The purpose of this section is to help the reader understand that in the commercial world the execution of MR projects and the acquisition of information relevant to decision making does not always proceed in a systematic step-by-step manner. A few factors that contribute to compression, adaptation, and modification of the MRP discussed previously are listed below:

- Information is not just acquired through formal MR studies, but it is also acquired

through secondary and syndicated data sources.
- Time compression – fire fighting is a common feature of daily life at companies.
- Demonstration effect – management ideas are not exempt from fashion cycles.
- *Inertia* – in the case of MR, defined as continuous motion – keeping a study going, as opposed to launching it *ab initio*.

A brief discussion of how each of the above factors modifies the traditional-rational MRP follows.

Secondary/Syndicated data. Several companies, like Nielsen, IRI, Corporate Executive Board, Forrester, and Gartner are in the business of producing and selling information. Their clients typically include companies like P&G, Hershey, HP, Citi, and Pfizer – companies with large flourishing MR departments who routinely use MR information and insights to inform their decision making. In short, companies use a mix of primary research (research that they initiate and conduct themselves, or through their agencies) and secondary/syndicated research (research designed and conducted by other companies) to build their market information stores.

In the case of secondary/syndicated data, the MRP described in Figure 1 may not be followed literally. Stages may be either compressed or skipped. Take the instance of Nielsen or IMS, two of the largest aggregators of market data, the former aggregating retail sales data and the latter physician prescription data. Once a company decides that it will be aggregating certain types of data, the sequence of decisions may shift very rapidly to critical variables like sample design, data recording, and data base creation.

From a company's point of view as well, the goal is different. Take the instance of P&G and Nielsen. On 22 July 2008, the Nielsen Company announced that it had extended its agreement with The Procter & Gamble Company to provide US consumer marketing information and insights. Under the multiyear agreement, Nielsen will provide P&G with marketing information and technology services, including retail tracking, insights on consumer purchase behavior, marketing analytics, and Nielsen Answers™, Nielsen's technology-based business intelligence solution. Nielsen will help P&G understand and develop new ways to market to its consumers.

What was P&G's motivation to extend the contract with Nielsen? Not some narrowly defined business or marketing problem, but a broadly defined business need. In the words of Garry Terrell, section manager, U.S. Market Measurements, P&G – "*the collaboration with Nielsen will enable us to access the data and research we need to drive business-building insights and innovation that both excites and delights our U.S. customers.*"

Time compression. In certain cases, for reasons associated with fire-fighting drills and/or crisis management, there is not enough time for full-blown formality and a detailed consideration of all decision variables in a systematic, sequential manner. Take the instance of pharmaceutical marketing. The Food and Drug Administration (FDA) and pharmaceutical manufacturers often use "*Dear Doctor*" letters to alert physicians about drug safety. For example, in March 2004, the FDA asked all manufacturers of a typical antipsychotic medication to add a warning statement describing the increased risk of hyperglycemia and diabetes in patients taking these medications. Drugs in the antipsychotic class covered by the FDA request included Zyprexa® (olanzapine, Lilly) Clozaril® (clozapine, Novartis), Risperdal® (risperidone, Janssen), Seroquel® (quetiapine, AstraZeneca), Geodon® (ziprasidone, Pfizer), and Abilify® (aripiprazole, Bristol–Myers Squibb and Otsuka American Pharmaceutical).

Dear Doctor letters do have a negative impact on physician prescribing. To what extent and for how long, is the key unknown. It is not unreasonable to expect marketing managers at all of the companies mentioned above to scramble to launch studies to get more data on those unknowns for their respective businesses. Since speed is of the essence, the following deviations may take place from the MRP outlined in Figure 1:

- A single research objective, driven by external circumstances.

- Opting for a research design that can deliver results faster, even though it may not be optimum.
- The need for speedy data collection may completely determine the types of questions and the type of data collected.
- Analysis may be sacrificed – reporting of counts and frequencies may suffice.

Demonstration effect. We live in a networked world. News of companies trying new and interesting things are reported frequently. In addition, company representatives are constantly showcasing their achievements and competencies at conferences. It is not unusual for an executive from a certain company hear about what a competitor or another admired company has done and want that replicated in his/her own company.

The use of ethnography in MR provides an excellent example to understand how the demonstration effect can shrink the formal MR process. In recent years, there has been a sharp increase in the use of ethnographic approaches in qualitative MR to gain a deeper and more context-bound understanding of customers, their attitudes, and their behaviors. Popular media, like BusinessWeek, Wall Street Journal, and Fortune have been quick to catch on and have featured full-length features on how the most innovative companies in the world are hiring anthropologists and using ethnography to boost their innovation and product design efforts.

To illustrate, the 14 March 2007 issue of BusinessWeek carried an interview with Jan Chipchase, an interface designer and a user anthropologist at the Nokia Research Center. The article reported how ethnographic research approaches were used to understand about how people share objects, such as cell phones. In Africa, and several other places, increasingly a cell phone is shared. A family might have one. A village might have one, or someone who runs a phone kiosk in a village might have one. The purpose motivating the research was idea generation concerning how Nokia could redesign the mobile phone and the communication experience to be more suitable for sharing.

An executive at HP, Intel, or some other PC company may read this article and ask, if Nokia can do this why can't we? After all Internet cafes where computers are shared already exist in several parts of Asia and Europe. The appeal of a new idea tried by a competitor, or an admired company, may be strong enough to circumvent much of the formal MRP. The brief may be direct and simple – to conduct ethnographic research in less affluent countries to understand PC sharing behavior and the demands that it places on machine and user interface design.

Inertia. The fundamental assumption underlying the traditional MRP is that a study is to be launched. But in several cases, referred to as *longitudinal research*, *continuous research*, or *tracking research*, the study has already been launched in some previous time period. The goal of the researcher therefore is not to launch a study, but to sustain its inertia – keep it going.

Examples of continuous or tracking research from several different sectors are presented below:

- Technology companies like IBM and HP may conduct tracking research to understand purchase and consumption trends for hardware, services, printers, and consumables, like printer ink.
- Service companies like restaurants, banks, and retail stores conduct customer satisfaction tracking studies to monitor customers' experience with the store, its offerings, and its front-line customer facing employees.
- Automobile and consumer electronics companies conduct new buyer tracking research to monitor buyer experience and postpurchase dissonance or regret.
- Social research companies conduct tracking research to monitor shifts in attitudes toward public policy (education, healthcare, trade, subsidies, taxation); political polling companies conduct tracking research to monitor shifts in preferences for different candidates contesting an election.

Since the research is already in motion and the principal goal is to provide an unbroken basis for making comparisons and drawing conclusions, the fundamental focus of the MRP is

housekeeping. Barring occasional adjustments to sample design, question wording, and measurement scales, the principal focus is on flawless and timely implementation. Very rarely does the researcher cycle through the entire MRP.

Marketing Research and Its Discontents: Implications for the MRP

Every discipline has its own grammar and MR is no exception. Traditional MR is predicated on a simple rule – a business issue can be translated into a set of questions, qualitative or quantitative, which can then be asked to respondents, who in turn want to and are able to answer them, which responses when properly analyzed result in meaningful insights.

In most of its early years, this formula worked very well. Companies that embraced and adopted emerging principles and techniques of MR and used its output intelligently were actually able to generate a competitive advantage. Reports of business success using specific MR techniques spread quickly with the help of academic literature and the business press, creating fashion cycles and mass adoption. Attitude research using Likert scales, advertising effectiveness research based on low involvement theories, trade-off analysis using conjoint research, and customer satisfaction research to name a few have been used enthusiastically and extensively. And while there was the occasional highly publicized howler, perhaps none more famous than the introduction and subsequent withdrawal of New Coke, success stories – like the design of the Marriott Courtyard, the introduction of Ford Taurus, and the use of segmentation research by banks and credit card issuers – continued raking up points for MR.

In recent years, the relationship between MR and its key stakeholders has soured, the extent of which varies depending on who is providing the reasons. As with most things that are successful, the expectations from MR have increased tremendously. And the general opinion is that MR has not responded effectively. The reader might wonder why we are discussing the practice of MR, when the focus of the article is on the MRP. The reason is simple. How MR is designed and executed (MRP) is a direct function of how MR is conceptualized and used.

It would be difficult, and its not necessary, to document every grievance against MR. As such, only a few important criticisms are recognized below.

- MR is very transaction oriented, one project at a time, and has never really evolved into a true information or a decision support system.
- The function and the process are very technique driven – greater focus on research designs and analytics than on business issues.
- Businesses today face greater uncertainty and risk, emphasizing anticipation and prediction – traditional MR tools are more suited for description.
- Asking consumers questions is only one source of information and insights.
- Managers coordinate business systems to drive business performance – like profitable long-term growth; MR should focus on the entire system and not just elements of it.

While preparing to write this article, the author had informal in-person and telephone discussions with several executives (approximately seven) in charge of MR departments at large product/service companies in the Financial, technology, consumer packaged goods, and healthcare sectors. The purpose of the conversation in each case was to understand how the MRP actually unfolds in their companies. It is important to point out that these companies have made significant changes in the way they have organized their businesses and the MR function, including what they call it. The term *MR Department* has all but vanished; today most companies prefer labels like Customer Insights, Consumer Insights, Market Intelligence, and so on.

Key points that emerged in the discussions are presented below.

- Unlike the traditional MRP, the starting point is not a business decision, but a business agenda. The focus is not so much on a single decision maker, but on a collection of decision makers of a larger unit – such as a category (beverages), segment (on-the-move consumption), or market (retail banking).
- There is a greater emphasis today to plug information gaps, not merely fulfill research

obligations by checking off research objectives.
- There is greater pressure on MR to adopt a comprehensive, multidimensional, approach to information; data integration is at a premium.
- Distinct preference for influencing business outcomes, not just generating results and insights.
- Greater expectations of the function to be effective along a continuum ranging from business maintenance to business transformation.

This is not to say that the traditional MRP is dead and will never be employed. On the contrary, for projects that can be structured and where there is a dominant single need – to test a concept, to discover subsegments in a growing market, to estimate price and promotion elasticity – traditional MRP still is king! However, the number of situations that fit the above description is decreasing, owing to the fundamental changes in how businesses create and deliver value to their customers. Consequently, traditional MRP will continue to contribute, albeit to a smaller fraction of the total information needs of a company or a business unit.

As mentioned in the previous paragraph, significant changes are taking place in the environments companies operate in today and in the way they conduct their business. The most important among them are as follows:

- Greater environmental uncertainty and complexity.
- Future and organic business growth both likely to be extremely challenging for virtually all companies.
- Increase in social media usage, user generated content, and information generation in peer-to-peer forums and aggregator forums, most of which is free.
- Increased collaboration and interdependence between companies and all stakeholders in a company's ecosystem – other companies, governments, nongovernment institutions, customers, consumers, and suppliers.

These changes affect how companies gather and use information. They also influence the perceptions of the companies in terms of what information outcomes they consider to be useful and relevant. This article would be doing a disservice to the reader if it did not draw his/her attention to these developments and the implications it has for the practice of MR and the implementation of information acquisition and dissemination programs.

On the basis of discussions with a small set of practitioners and the developments discussed above, a forward-looking perspective is provided in Figure 4. It is labeled – A proposed information process for influencing business outcomes – to accurately depict its focus; research data obtained by asking questions to respondents is only a part of this more comprehensive focus. The proposed model is conceptualized to have four macrophases as follows:

- *Issue identification*. The purpose of this phase is to generate an inventory of issues relevant for the achievement of current business goals and for helping bring to market future business aspirations
- *Information need assessment*. The focus of this phase is to actually start at the end, articulate key information that will be needed for the system to implement initiatives to achieve its business goals, and then work back to identify information gaps
- *Integration of insights and perspectives*. From a large number of sources (open source, secondary, syndicated, primary); using several types of data (self reported, web conversations, opinions and ratings, attitudes, behavior); using a variety of data collection methods (obtrusive, nonobtrusive, observational, surveys); and a variety of analytical tools (analysis, mining, extrapolation prediction, modeling)
- *Inspire execution*. The next generation information system must inspire and guide execution and discuss implementation alternatives with accompanying risks and benefits.

Conclusion

Companies like, P&G, Capital One, Hershey, and HP regularly spend large amounts of resources, human and financial, on MR projects to improve the quality and effectiveness of their

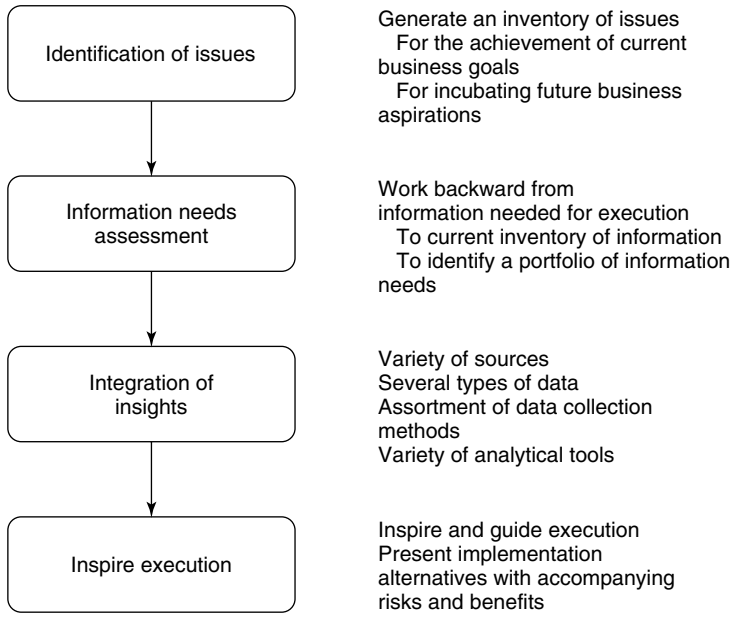


Figure 4 A proposed information process for influencing business outcomes.

business and marketing decision-making. As a large number of professionals and agencies are involved in conducting MR studies, both clients sponsoring MR projects and companies offering MR services stand to benefit if standardized processes can be followed by all involved. The MRP is a key component to standardizing the approach to data collection and ensuring a consistent, acceptable quality of research output. It is for this reason that the article discusses the MRP, its macrophases and the key decisions in each phase in detail.

While the need for uniformly high quality of research output is ubiquitous, the operating conditions in companies on a day-to-day basis are rarely consistent or homogeneous. Consequently, the article also discusses conditions under which researchers modify and adapt the traditional MRP. Additionally, while MR has made admirable contributions to business decision-making, demands for MR to morph into a larger and more comprehensive information system for influencing business outcomes are increasing. Accordingly, the article presents a prototype that has the potential of meeting these emerging expectations. The article acknowledges that the traditional MRP will continue to play a vital role. But it also urges for a bigger playground where other forms of data, data collection methodologies, and analytical procedures feature equally prominently, sometimes even more.

Bibliography

Bonabeau, E. (2002) Agent-based modeling: methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 7280–7287.

Clandinin, D.J. and Connelly M.F. (1989) Narrative and Story in Practice and Research. EDRS Confirmit Software, Oslo, www.confirmit.com.

Corporate Executive Board www.executiveboard.com. Washington, DC.

Forrester www.forrester.com. Cambridge, MA.

Gartner www.gartner.com. Stamford, CT.

Kumar, V., Aaker, D.A., and Day, G.S. (2009) *Essentials of Marketing Research*, John Wiley & Sons, Inc., New York.

Malhotra, N.K. (2007) *Marketing Research: An Applied Orientation*, Prentice Hall, Upper Saddle River, New Jersey.

MarketTools www.markettools.com. San Francisco, CA.
Nielsen www.nielsen.com. New York.
SAS www.sas.com. Cary, NC.
SPSS www.spss.com. Chicago, IL.
Surveywriter www.surveywriter.com. Chicago, IL.
Zikmund, W.G. and Babin, B.J. (2007) *Exploring Marketing Research*, South Western-Cengage, Florence.

motivation research

Robert V. Kozinets

Questions such as "Where did Betty Crocker's maternal appearance come from" or "How was the famous Exxon logo 'Put a tiger in your tank?' coined?" or "How did Nestle tried to crack the cultural code and market coffee to the tea-drinking nation of Japan?" or "How did ice cream containers become round?" have answers that are quite surprising, as they are all involved with a very unusual and controversial form of consumer research called *motivation research*, which continues to be influential and contentious to this day.

Motivation research is a term used to refer to a selection of qualitative research methods designed to probe consumers' minds to discover the deep, often subconscious or latent reasons and goals underlying everyday consumption and purchasing behaviors. Motivation research was the premier consumer research method used in the 1950s, leading to its lasting influence on the areas of advertising and consumer research, as well as on advertising practice. In academic circles, however, the rapid rise of motivation research was followed by an equally rapid decline (Stern, 2004).

Sigmund Freud's psychoanalytic personality theories provide the foundation for the development of motivation research in marketing. The theory was built on the premise that unconscious or deeply hidden needs and drives underlie all human behavior. Particularly influential are sexual drives and other deep-seated biological ones such as the need for dominance and aggression. Freud's theory was constructed from his own patients' recollections of early childhood experiences, combined with an analysis of their dreams, as well as the specific nature of their problems with mental and physical adjustments to their situations. In Freudian theory, these needs are assumed to lie at the very core of human personality and motivation.

Ernest Dichter was trained as a psychoanalyst in Vienna and moved to New York in 1938 (Stern, 2004). He began adopting Freud's psychoanalytic techniques to the study of consumers buying habits in the 1930s. Until that time, marketing research has mainly focused on what consumers did – descriptions of their behaviors (*see* DESCRIPTIVE RESEARCH). These studies tended to be quantitative as well as anecdotal, what came to be commonly known as *descriptive statistical case studies*. Providing a fresh and refreshing change to this, Dichter used qualitative research methods not to describe what was actually done, or even said; instead, he tried to delve deeper into why consumers actually behaved as they did (*see* EXPLORATORY RESEARCH).

Dichter's fame quickly spread because of his new approach, engaging presentations, and skilled writing style. The entertaining and usually surprising explanations offered for different kinds of consumer behavior fascinated the advertising and marketing world, particularly since many of these explanations were grounded in sexual motivations. For example, marketers were told that cigarettes and lifesaver candies were bought because of their sexual symbolism. Men, it was claimed, regarded their convertible cars as surrogate mistresses (Dichter, 1964). Before long, almost every major advertising agency in the country had on its staff a psychologist who conducted motivation research studies. Some of Dichter's more influential and long-lasting insights can help us understand the basic principles of the analysis.

In one set of studies, Dichter describe baking as an essential expression of femininity as well as motherhood. It was also very sensorily fulfilling, evoking nostalgic memories such as the delicious odors that pervaded the house when the mother was baking. Dichter (1964) postulated that, when baking a cake, a woman symbolically linked to the act of giving birth. The most powerful moment of the experience occurs when a product is actually pulled out from the oven, symbolizing the moment of birth. When a woman bakes a cake for a man she is offering it to him as a symbol of her fertility. At the time, the envisioning of Betty Crocker's appearance was apparently based on Dichter's analysis of baking as fertility. In our contemporary times, we can only speculate what is symbolized by the baking of a cake by a husband for his wife (Hitt, 2000).

The amazingly long-lived Exxon/Esso logo "Put a tiger in your tank" and its related campaigns had their genesis in another Ernest

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Dichter study (Dichter, 1964). According to Dichter, the automobile allows consumers to convert into reality two powerful subconscious urges. The first is the aggressive urge to compete with, beat, best, and destroy that psychoanalysis finds present as a real force in all human psyches. The second is *Thanatos*, the powerful encounter with mortality, the fear of death. For example the expression "step on it" comes from the desire to feel power. The phrase "I missed hitting that car (or baby stroller) by inches" reflects the profound desire to play with danger, to come close to the edge of mortality. Dichter's idea was to develop a slogan that allowed the fuel company to tap into consumers' unconscious aggressive motives for driving a car (Patton, 2002). The outcome emerging from this research insight, added to the creativity of Exxon's hired advertising copywriters, was an image that mixed control of a wild force with a tamed viciousness: a tank full of striped, feline fury.

And how did ice cream boxes become round? After an analysis conducted for a dairy food company, Dichter (1964) found that ice cream was a type of sexually satisfying, orgiastic food. It melted in one's mouth providing a extremely sensual and pleasurable sensation. Like mother's milk, it did not need to be chewed, but invited sucking. It was sweet and creamy, a pure hedonistic pleasure. Its sweetness and richness signify incredible abundance and people eat it as if they wanted it to run down their chins. Because of these aspects of pure abandon and hedonistic excess, Dichter recommended an ancient symbol of limitlessness. The circle, the line without end, the ceaseless pleasure of abundance and earthly delights, was the best packaging for ice cream. Further, the box should have illustrations running around its periphery to show the unending delight of the delicious ice cream within.

This form of analysis may seem foreign and unscientific, and some of these examples might seem overblown and hyperbolic. The fact that motivation research seems strange to us really is not very strange at all. It is based on the premise that we, like all consumers, are not usually aware of the reasons that underlie our behaviors (*see* ETHNOGRAPIC RESEARCH; EXPLORATORY RESEARCH). One of the reasons we react with embarrassment when we have these explanations is because they deal with such central, deep-seated needs. As motivation research digs deep into these needs, we gain insights that allow the marketer to better understand the underlying feelings, attitudes, and emotions that concern the use of a product, service, or a brand, or other consumption goods such as experiences and ideas. Perhaps it is therefore no wonder that, although she has had a little cosmetic work done, Betty Crocker is still Betty Crocker, Exxon still puts tigers in our fuel tanks (and in their advertising campaigns), and ice cream containers are still round. As seen later in this article, motivation research, repackaged into contemporary forms, continues to exert a major influence on advertising and marketing practice.

Evaluations of Motivation Research

By the early 1960s, marketers began to believe that motivation research was flawed (Stern, 2004). First, because of its intensive nature, qualitative research sample sizes had to be small. Traditional, statistics-driven marketing researchers worried about generalizing the findings of a small group of consumers to the totality of the marketplace (*see* STATISTICAL APPROACHES TO DETERMINING SAMPLE SIZES). Second, marketing researchers also worried that the analyses of projective test and depth interviews were highly subjective (*see* FOCUS GROUPS AND DEPTH INTERVIEWS; PROJECTIVE TECHNIQUES). In fact, it seemed that the quality of the interpretation depended very much on the acuity and even brilliance of the interpretive researcher (however, the same might be true of any technique, such as the detailed analysis and interpretation of sophisticated quantitative data).

Third, critics of motivation research noted that many of the projective tests that were used had originally been developed for purposes of clinical diagnosis of mental illness rather than for studies of marketing or for consumer behavior (*see* PROJECTIVE TECHNIQUES). Consumer behavior studies, however, were interested in finding explanations for the behavior of typical consumers. On the other hand, Freudian theory was developed in an entirely different social

context – nineteenth century Vienna – and applied in the 1950s into postwar America.

Finally, many motivation researchers would inject highly exotic and usually sexual reasons into what seemed to many to be rather prosaic and everyday consumer purchases and behaviors. For example, was it better to sell a man a pair of suspenders as means for holding up his pants, or as a type of protection to help him cope with his castration anxiety? Should a car be sold to a woman as an efficient vehicle or as an impressive substitute penis to help her compete in a world of aggressive men? Marketing researchers concerned with the dignity, reputation, and legitimacy of their scientific field began to question the often rather spicy explanations that were offered by the motivation researchers. It seemed that in the eyes of the motivation researcher, almost any mundane product or service also had a profoundly symbolic – and usually sexual – side. But for many marketing and consumer researchers, it was beginning to seem as though motivation research had forgotten Freud's famous saying that, sometimes, a cigar was just a cigar.

Motivation Research Today

Despite its rocky history and many critiques, motivation research is still regarded as an important technique by marketers who want to gain a deeper understanding into why consumers act in the ways that they do. These insights are often thought to be much more revealing than the information provided by traditional descriptive and quantitative marketing research methods (*see* FOCUS GROUPS AND DEPTH INTERVIEWS). Although the term *motivation research* is most often used to refer to qualitative research that is designed specifically to discover consumers' hidden, tacit, latent, or unconscious motivations, the term can also be used to reference any form of research that seeks to explain why people do things rather than simply describe what they do, or offer correlates to that behavior (*see* EXPLORATORY RESEARCH).

Over time, we have learned that many of the motivations driving consumer purchases indeed are motivated by sexual, dominant, violent, other biologically basic needs identified by Freud. We need look no further than the worlds of fashion, sports, entertainment, pornography, and video-game production and marketing for examples. The recent boom in design and attention to user-oriented design principles draw our attention to the fact that human motivations to buy and use products and services are indeed complex and social. In fact, because motivation research can often reveal unsuspected consumer motivations that underlie the use of a particular product or brand, one of its main uses in marketing is in innovation. Today, motivation research is popularly employed as a source of valuable, creative, front-end insights that are later quantitatively tested on larger representative samples of consumers. In a competitive global environment in which insight and innovation are keys to success, it is no wonder that motivation research continues to play an extremely important role.

Motivational research has a long, proven track record of being able to help deliver ideas that spawn new products and market categories, to act as a valuable input into the brainstorming of new products and services, to help in repositioning existing brands, as well as to develop new ideas for promotional campaigns. Motivation research offers novel ideas that can be used to penetrate the consumer's conscious awareness by appealing to their needs, fears, dreams, and desires that lie hidden under the surface of their conscious awareness.

By the 1980s, qualitative consumer and marketing research had evolved to encompass a variety of research approaches, including ethnography (*see* ETHNOGRAPIC RESEARCH), semiotics, CONTENT ANALYSIS, literary techniques, historical methods, discourse analysis, phenomenological methods, and conversation analysis (*see* EXPLORATORY RESEARCH). For some time, the two techniques of depth interview (*see* FOCUS GROUPS AND DEPTH INTERVIEWS) and projective tests (*see* PROJECTIVE TECHNIQUES) were the methodologies most commonly associated with motivation research. The psychoanalytic interview developed in marketing and consumer research, as it did in other social scientific fields, into the "depth" or (later) "long" interview, and the clinician's Rorschach and thematic tests evolved into a variety of projective tasks.

By the 1990s, many of these research approaches had become firmly established in marketing and consumer research. Motivation research can be seen as one step in a series of stages in the development of a vast range of marketing research techniques designed to uncover various facets of people's thoughts, desires, and lifestyles. Although projective tasks are used by researchers today, they are not nearly as common as they were during the heyday of motivation research. Interviews, however, are a staple of consumer and marketing research, whether in their academic one-on-one format, or in the focus group interview format, which is extremely popular in marketing research practice (*see* FOCUS GROUPS AND DEPTH INTERVIEWS). In the first years of the twenty-first century, these methods, which had their inception in motivation research, continued to make gains in enhancing our understanding of consumers' deeper motivations, meanings, and understandings of the world. Through its more detailed, complex, and subtle social analyses, motivation research is alive and well and still very much with us. By emphasizing the less obvious aspects of consumer behavior, it has helped marketing recognize some of the important forces driving consumer culture.

Some of the qualitative research techniques used in the original motivation research have been developed and are now in common use, particularly among advertising, marketing, and marketing research practitioners. In the following section, we present an introduction to some of these tools that may be useful in research.

Motivation Research Tools and Techniques

In contemporary marketing research, there are a number of qualitative research techniques used to delve into customers' unconscious hidden motivations. These techniques draw their lineage directly from motivation research. They include projective tests and tasks, metaphor analysis, storytelling, word-association tasks, sentence-completion tasks, picture generation, and also photo sorts (*see* EXPLORATORY RESEARCH; FOCUS GROUPS AND DEPTH INTERVIEWS; OBSERVATION METHODS; PERSONAL OBSERVATION; PROJECTIVE TECHNIQUES). Marketing researchers such as Gerald Zaltman and Clotaire Rapaille are new practitioners of the craft of motivation research who pick up, develop, and make contemporary the approach initially popularized by Ernest Dichter.

Thematic apperception test and other projectives. The thematic apperception test was developed by Henry A. Murray and Christiana D. Morgan at Harvard University during the 1930s to explore some of the motivating forces underlying personality. It is a picture interpretation technique that uses a standard series of 30 pictures about which the research subject is asked to tell a dramatic story. The pictures themselves are open-ended yet provocative, showing, for example, a young boy contemplating a violin that sits on a table before him.

For a marketing research example, consider the hypothetical research we might conduct for a new cold medication. In this research, there would be an image of a middle-aged woman in a bathrobe looking into a mirror and a caption underneath it saying "She looked into the mirror that morning and realized that she had a cold." When we show this picture to women, what might we hear? As researchers we would probably discover that middle-aged women consider their lives to be fast paced, full of obligations, and socially active. The discovery of a cold abruptly disturbs the pace of their lives. An advertisement that could result from this projective research might show an attractive, busy, middle-aged woman walking down a busy street. She sneezes and then, suddenly, all the motion around her stops. She opens her purse, takes out the cold medication, takes it, smiles, and then her life speeds up to its previous rapid pace. By providing such an image in advertising, the marketer is to symbolize that they understand the lived experience and social world of the consumer. This image-based positioning builds strong emotional ties between the brand and the consumer.

Metaphor analysis. Metaphor analysis is a method based upon providing visual images to consumers, or having them collect their

own images, and then using these images for projective tasks in which they compare and contrast products, services, or brands to the various images, and the images to one another. The analytic goal of metaphor analysis is to generate guiding analogies providing a deep sense of understanding of how consumers relate to a particular product, category, or brand.

Gerald Zaltman, Professor Emeritus at the Harvard Business School, has popularized this method in recent years. His approach, which he terms ZMET, or the Zaltman metaphor elicitation technique, is based upon many of the same fundamentals as classic motivation research. Zaltman (1996) offers the following founding principles for his metaphor test: the nonverbal and unconscious nature of most social communications, the image-based or imaginary nature of most human thought, the centrality of the form of metaphor to human thinking, the embodiment or bodily nature of cognition and thought, the linkage of reason, emotion, and experience and the assumption that deep structures of human motivation can be accessed through projective-style research (*see* PROJECTIVE TECHNIQUES). The relation of the body, emotions, and hidden unconscious motivations to the projective task of marketing research relates this work directly to motivation research.

In the ZMET test, consumers are either asked to collect a number of images that represent their thoughts and feelings about a product, service, or brand, or they are provided with these images in various forms. The forms can be clippings from a magazine or specially designed graphics and photographs that they select from the computer screen. Research participants are given the focus of the research a week or more in advance of the actual interview. They are encouraged to ruminate and think deeply about the topic of the assignment and, if relevant, about their selection of photographs and images. The research question must be carefully considered in order to focus the response of the consumers. Considerations may include their opinions of a particular company or brand, their experiences of a purchase setting or a buying process, their use of a service or product, or how they feel about a certain concept. Participants are also directed not to choose photographs that literally represent a product or service. For example, if the research topic is mobile phones, the research participant would be asked not to select any actual photographs or graphics of mobile phones. Forcing the research participant to choose photos that indirectly relate to digital cameras activates an analogical style of reasoning that, it is assumed, can help reveal latent feelings, thoughts, and motivations (Zaltman, 1996).

The interview is approximately two hours in length and consists of guided questions regarding the images chosen by the consumers to answer the focused research question. As the interview progresses, participants are asked for their opinions regarding other senses that might express their feelings and thoughts. The researcher's main task is to attempt to elicit as many rich metaphors from the participant's experiences and memories as possible. The idea, according to Zaltman (1996, p. 15), is to allow "deep, latent ideas to emerge as well as for the expression of the wide range of relevant ideas." Increasingly, the ZMET and other metaphor-based techniques have incorporated advanced mapping and diagram-drawing techniques to analyze and present the findings of their research, as well as digital graphical design and even creative animations with voice-overs. The ZMET technique has been very successfully adopted and used by the top corporations in the world, from Coca Cola and Procter & Gamble, to Walt Disney, Mercedes-Benz, Bank of America, Microsoft, and Chevron.

Storytelling. In the method of guided storytelling, consumers tell real-life stories about the meanings or uses of products under investigation. Although this technique is related to depth interviews, it is considered to be a distinct form of research. Often, the storytelling method asks research participants to imagine and relate stories relating to their own product or service usage. Another application of the storytelling methodology requires subjects to imagine a story involving another person. So, for example, people who have a fear of tall buildings would be asked to imagine and then tell a story about why some people are afraid of tall buildings. In this way, the storytelling process ameliorates people's own anxiety, embarrassment, and social-censoring mechanisms. Doing so, people

will be less likely to censor their own apprehension about heights and to offer an accurate portrayal (*see* EXPLORATORY RESEARCH).

As a form of motivation research, the storytelling method seeks to plumb the depths of consumer motivations. For example, Kimberly Clarke used a storytelling method in order to study current perceptions about diapers (Lieber, 1997). They found that using this research method, the parents actually considered that diapers were a type of clothing related to a particular stage in the child's development. If their child wore diapers for too long, the result was that the parents became distressed and embarrassed because they viewed it as a failure. They felt that they had not toilet trained their children properly and that this lack of success was obvious from the children wearing the wrong apparel for that particular stage in their life. Using the data from this storytelling study, Kimberly-Clark introduced its new Huggies Pull-ups training pants. These training pants introduced a highly successful new category into the US diaper industry.

Word-association and sentence-completion tasks. In word-association tests, research participants are presented with words one at a time, and then asked to respond with the first word that comes to mind, for example, "What is the first word that you think of when I mention the word or category coffee?" In a sentence-completion task, respondents are asked to complete a sentence upon hearing the opening phrase. For example "People who drink Starbucks are . . ." or "A Starbucks latte reminds me of. . ." These related methods can be very useful in determining consumers' associations with existing brand names, eliciting related choice sets, as well as determining associations with new brand names that are being considered and that are currently under development.

An entire web site has been devoted to creating word clouds based on people's reactions to brand names. It is called *Brand Tags* and is available at http://www.brandtags.net. As can be readily seen, by clicking on almost any brand name, there is a very wide assortment of responses. For example, the word cloud associated with Starbucks includes not only strong, burnt, green, and trendy but corporate, mermaid, addictive, ubiquitous, and overpriced.

Picture generation. Marketing researchers can also use visual images to study consumers' perceptions of various brands. They can use an analysis of consumers' guided drawings or doodles to help understand consumer perceptions, and use that understanding to brainstorm new strategies for advertising. Consider some hypothetical research on coffee. Research participants were asked to draw pictures of the typical drinker of Maxwell House coffee. These drawings might elicit drawings of old-fashioned, chubby females wearing frilly aprons. When asked to draw pictures of Starbucks drinkers, the drawings might show a series of slim, cool, "with it" women wearing high heels and miniskirts. For a company like Maxwell House, these findings might provide important input about the dire need to reposition its product to seem more in tune with the times.

Photo sorts. In photo sorting tasks, respondents receive stacks of photos depicting various events and are asked to select the pictures from the set that best portrays or captures some particular element that the researcher is interested in investigating. In a photo-sort study that was conducted by an advertising agency for Playtex, the manufacturer of bras, research participants received a stack of photos that portrayed many different types of women wearing many different types of clothing. First, the research participants were asked to choose pictures that represented the typical user of Playtex bras. They chose overweight, old-fashioned, big-breasted women (O'Shaughnessy, 1995, p. 437). These women, who were Playtex users themselves, were then asked by the researcher to select the pictures from those that best captured their own self images. Although many of the respondents may have been overweight, full-breasted, and old-fashioned in appearance, they selected photos that showed physically fit, well-dressed, and independent-looking women. The advertising agency then advised Playtex to stop stressing the comfort of its bras in its advertising campaigns and, instead, to design a new campaign that showed thinner and sexier

big-bosomed women under the slogan "the fit that makes the fashion."

Contemporary Motivation Researcher: Clotaire Rapaille

One of the great contemporary practitioners of an updated form of motivation research is Clotaire Rapaille, a French-born psychoanalyst and medical anthropologist whose methods for mining covert motivations have compelled many Fortune 500 companies to spend massive amounts of money on his brand of "culture code" marketing research. Rapaille's research is based upon a sociobiological idea of "imprinting" that originated with studies of geese:

> the combination of experience and its accompanying emotion create something widely known as an imprint, a term first applied by Konrad Lorenz. Once an imprint occurs, it strongly conditions our thought processes and shapes our future actions. Each imprint helps makes us more of who we are. The combination of imprints defines us. . . every imprint influences us on an unconscious level. Rapaille, 2006, pp. 6–7.

To explain how he uses the notion of imprinting in his research, Rapaille gives the example of working with Swiss food giant Nestlé trying to sell instant coffee in Japan. He discussed how he gathered groups of people together to discover how they imprinted their perceptions of coffee. He scheduled a 3 hour session with each group of research participants. In the first hour, he queried the participants about coffee usage as if he was an alien from a different planet who knew absolutely nothing about coffee. In the second hour, he had them sit on the floor using scissors and piles of magazines to make a collage of words that represented their impressions of coffee. His intuition was that these visual images would help them to express stories about coffee that could give him further insights into coffee consumption.

For the final hour of his investigative research, Rapaille asked participants to lie on the floor with pillows, put on soothing music, and asked them to relax. He spoke to them about the past, taking them from their adult years to their teen years to a time when they were young. At that point in the guided meditation, he asked them to think about coffee and to recall their earliest memory of it. He sought to elicit the very first time that they consciously experienced coffee as well as their most significant memory of coffee. Rapaille says that he learned that Japanese people had only a very superficial imprint of coffee and many had no imprint of it all (Rapaille, 2006).

His conclusion to Nestlé was that the current strategy of attempting to switch to drinkers to coffee was bound to be a failure. Nestlé instead needed to begin by giving coffee a meaning, in other words, offering up an imprint for the Japanese market (Rapaille, 2006). Nestlé began acting on this advice by creating desserts for children with coffee flavors but without caffeine. Because they were gaining a sweet-tasting coffee experience as their first imprint, Japanese youth gained a very positive first impression. The idea was that this impression and imprinting would follow the youth through their lives. Although cause and effect are difficult to measure in such complex cultural circumstances, coffee sales in Japan have steadily risen and now approach up to a billion pounds per year.

Betraying his psychoanalytic roots, and, of course, in the world of marketing research, the path from Freud to Dichter being rather short, Rapaille considers that he reveals the "reptilian" part of the brain, which is the home of violence, strong smells, violence, sex, and primal emotions (Hitt, 2000; Rapaille, 2006). According to Hitt (2000), Rapaille typically begins a marketing research session by leading a group of about 20 people through a series of word-association games. After writing words on a board, the researcher then asks the group to identify themes that unite the words. He then has them tell stories based on the concepts that have been written on the board. The idea behind this technique is to generate a number of little stories. And as in the concluding part of Rapaille's (2006) investigative research, the marketing research session involves having the research participants lie on the floor, with blankets and pillows, while repetitive and relaxing music plays for about 20 minutes to calm down the active mind. At this point, the researcher seeks to take the room full of participants back to their earliest memory of a product or category. After talking to them, he asks them to write down the story of their earliest and their most vivid memories relating to

the product or category. These recollections then become the data that the researcher uses to find deep-seated, hidden, archetypal associations. Of course, we have no proof that these memories are actually valid memories of childhood experiences. Most likely, they are constructions that the participants rearrange in the present moment in order to fulfill the requirements of the research. Nonetheless, just like other types of impressionistic research gathered through motivation techniques – a collage, a completed sentence, a creative story told about an ambiguous picture – they become the basis of researcher insights in this method.

Reading through examples of Clotaire Rapaille's marketing research conclusions, one cannot help but be struck by its similarities to Dichter's motivation research. Rapaille considers that his research decodes the existing codes behind product categories discovering people's true product meanings and motivations for use. His colorful style, powerful and simple conclusions, creative metaphors, and usually biologically rooted bases for explanation evoke the spirit of motivation research. Consider as a final example, the way that Rapaille discusses the difference between the way that French and American people understand the category of cheese:

> The French code for cheese is ALIVE. This makes perfect sense when one considers how the French choose and store cheese. They go to a cheese shop and poke and prod the cheeses, smelling them to learn their ages. When they choose one, they take it home and store it at room temperature in a cloche (a bell-shaped cover with little holes to allow air in and keep insects out). The American code for cheese, on the other hand, is DEAD. Again, this makes sense in context. Americans "kill" their cheese through pasteurization (unpasteurized cheeses are not allowed into this country), select hunks of cheese that had been prewrapped – mummified, if you will – in plastic (like body bags), and store it, still wrapped airtight, in a morgue also known as a refrigerator. Rapaille, 2006, p. 25.

Bibliography

Dichter, E. (1964) *Handbook of Consumer Motivations*, McGraw-Hill Book Company, New York.

Hitt, J. (2000) Does the Smell of Coffee Brewing Remind You of your Mother? New York Times Magazine (May 7), **16**, p. 71.

Lieber, R.B. (1997) Storytelling: A New Way to Get Close to Your Customer. Fortune Magazine (Feb. 3).

O'Shaughnessy, J. (1995) *Competitive Marketing: A Strategic Approach*, 3rd edn, Routledge, New York.

Patton, P. (2002) Car Shrinks. Fortune Magazine (Mar. 18), p. 6.

Rapaille, C. (2006) *The Culture Code: An Ingenious Way to Understand Why People around the World by and Live As They Do*, Broadway Books, New York.

Stern, B.B. (2004) The importance of being Ernest: commemorating Dichter's contribution to advertising research. *Journal of Advertising Research*, **44** (2), 165–169.

Zaltman, G. (1996) Metaphorically speaking: new technique uses multidisciplinary ideas to improve qualitative research. *Marketing Research*, **8** (Summer), 13–20.

marketing research proposal

Gaurav Bhalla

Introduction

Imagine you are the Director of Brand Strategy and Communications at McDonald's corporation. You are aware that obesity, especially child obesity, is fast becoming a major health and social issue. You have a copy of the book *Fast Food Nation: The Dark Side of the All-American Meal* (2001) by investigative journalist Eric Schlosser, but have not read it from cover to cover. However you have seen the movie *Fast Food Nation*, based on the book. As a marketing and strategy professional, you know that McDonald's cannot ignore this growing trend. You and your colleagues firmly believe that the company should proactively signal to its numerous stakeholders – consumers, franchise owners, school teachers, nutritionists, media, legislators, and so on – that McDonald's is sensitive to the social and health implications of obesity and that it is genuinely interested in playing its part to help combat its growing menace.

However, while you are aware of the macro aspects of the problem, you lack a granular understanding of what is actually happening in the marketplace. You have numerous questions on the implications this growing concern about obesity has for McDonald's.

- What are stakeholders' specific attitudes toward the phenomenon of obesity and toward food companies like McDonald's, Kraft, Kentucky Fried Chicken, Hershey, Nestle, and so on?
- Do stakeholders believe that McDonald's is responsible for contributing to obesity? Is it a global perception, or is it localized in some countries?
- Do stakeholders apportion the blame on both the company and its menu offerings? If yes, is it possible to separate the two components?
- What are some of the behavioral implications of stakeholders' perceptions?
- Are there specific initiatives that McDonald's can undertake to neutralize negative-obesity-related perceptions that stakeholders may have against McDonald's?

Experience has taught you that rather than tackling these questions yourself it is best to call in the market(ing) research (MR) department and have them commission a program of formal information gathering to help fill the gaps in your understanding. So, you invite the MR lead assigned to your department for a meeting.

Now, let us switch our attention to the MR department. Contemporary management practice prides itself in data-driven decision making. One of the tools that companies use to gather and disseminate market data is formal MR. Accordingly, it is standard practice for companies like McDonald's, Kraft, P&G, Microsoft, Citi, Mercedes Benz, Tesco, and so on, to have formal MR, or customer/competitive/market insight departments. Traditionally, MR departments of client companies seldom execute MR projects, studies, or programs themselves. They call on specialist providers of MR services to do so. Table 1 contains a list of leading specialist MR providers, ranked by their most recent revenues.

It is here that we encounter the MR proposal (*see also* SURVEY RESEARCH). To begin the information-gathering activity, client companies usually issue a request for proposal (RFP). How many and which MR service providers are invited to submit a proposal varies from one client company to another (discussion on RFP practices is beyond the scope of this article). In its essential form, an RFP is an invitation to the specialist MR firms, requesting them to submit a detailed proposal, explaining how they would go about meeting the requirement of the RFP, such as data collection, sampling, timing, presentation of results, execution, and pricing. The RFP may be issued by MR professionals, those who have direct responsibility for coordinating the information-gathering activity with the MR firm, or by centralized sourcing and procurement departments, like the Purchasing Department. This article will assume that the RFP is issued by the MR department, allowing for greater flexibility of dialogue and presentation of ideas.

An MR proposal is not merely a compliance document, complying with all the requests and demands made in the client's RFP. It is also a sales document. A proposal is an opportunity for the MR firm to make a strong case for why it should be awarded the project. It is this

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Top 25 global research organizations. (Source: Marketing News, 2008.)

Rank 2006	Rank 2005	Organization	Headquarters	Parent country	Web site (WWW.)	No. of countries with subsidiaries/branch offices[a]	Research-only full-time employees[b]	Global research revenue[c] (U.S. $, in millions)	Percent change from 2005[d]	Revenue from outside home country (U.S. $, in millions)	Percent of global revenue from outside home country
1	1	The Nielsen Co.	New York	U.S.	nielsen/com	108	39,517	3696.0	26%	1726.0	46.7%
2	3	IMS Health Inc.	Norwalk, Conn.	U.S.	imshealth.com	76	7400	1958.6	8.9	1241.7	63.4
3	2	Taylor Nelson Sofres plc	London	U.K.	tns-global.com	75	14,570	1851.1	2.5	1572.6	85.0
4	5	The Kantar Group[e]	London & Fairfield, Conn.	U.K.	kantargroup.com	59	6900	1401.4	4.1	892.4	63.7
5	4	GfK AG	Nuremberg	Germany	gfk.com	57	7900	1397.3	5.4	1059.1	75.8
6	6	Ipsos Group SA	Paris	France	ipsos.com	50	6503	1077.0	6.5	952.5	88.4
7	8	Synovate	London	U.K.	synovate.com	52	5726	739.6	9.5	673.6	91.1
8	7	IRI	Chicago	U.S.	infores.com	8	3600	665.0	6.6	233.0	35.0
9	9	Westat Inc.	Rockville, Md.	U.S.	westat.com	1	1906	425.8	0.8	–	–
10	10	Arbitron Inc.	New York	U.S.	arbitron.com	2	1045	329.3	5.9	13.2	4.0
11	11	INTAGE Inc.[f]	Tokyo	Japan	intage.co.jp	2	1558	264.8	7.0	1.8	0.7
12	14	J.D. Power and Associates[e]	Westlake Village, Calif.	U.S.	jdpa.com	8	840	232.6	16.4	62.1	26.7

13	12	Harris Interactive Inc.	Rochester, N.Y.	U.S.	harrisinteractive.com	5	1106	216.8	2.6	47.8	22.0
14	13	Maritz Research	Fenton, Mo.	U.S.	maritzresearch.com	4	757	216.4	4.5	38.2	17.7
15	16	The NPD Group Inc.	Port Washington, N.Y.	U.S.	npd.com	13	950	186.9	13.1	41.3	22.1
16	15	Video Research LTD.[f]	Tokyo	Japan	videor.co.jp	3	398	173.7	0.2	0.2	0.1
17	17	Opinion Research Corp.	Princeton, N.J.	U.S.	opinionresearch.com	3	674	154.7	2.7	57.9	37.4
18	18	IBOPE Group	Sao Paulo	Brazil	ibope.com.br	16	1663	103.9	10.7	21.2	20.4
19	19	Lieberman Research Worldwide	LosAngeles	U.S.	Irwonline.com	4	298	78.3	5.7	15.2	19.4
20	–	Telephia Inc	San Francisco	U.S.	telephia.com	1	258	71.8	12.2	2.0	2.8
21	23	comScore Inc.	Reston, Va.	U.S.	comscore.com	5	377	66.3	31.8	5.7	8.6
22	20	Dentsu Research Inc.	Tokyo	Japan	dentsuresearch.co.jp	1	118	61.2	4.4	–	–
23	22	AbtAssociates Inc.	Cambridge, Mass.	U.S.	abtassociates.com	1	200	53.6	2.7	–	–
24	21	Nikkei Research Inc.	Tokyo	Japan	nikkei-r.co.jp	4	170	50.7	0.9	–	–
25	25	Burke Inc.	Cincinnati	U.S.	burke.com	1	205	50.0	16.8	6.5	13.0
					Total		104,639	15,522.8	5.1%	8664.0	55.8%

[a]Includes countries that have subsidiaries with an equity interest or branch offices, or both. [b]Includes some nonresearch employees. [c]Total revenue that includes nonresearch activities for some companies is significantly higher. This information is given in the individual company profiles. [d]Rate of growth from year to year has been adjusted so as not to include revenue gains or losses from acquisitions or divestitures. See company profiles for explanation. Rate of growth is based on home country currency and includes currency exchange effects. [e]Estimated by Top 25. [f]For fiscal year ending March 2007.

balancing act – balancing the technical aspects of how it will execute the project, with the selling aspect of why it should be awarded the project – that makes writing an MR proposal an art form. Accordingly, this article will discuss the MR proposal not as a recipe, but in terms of a framework that we turn to next.

Anatomy of an MR Proposal

Table 2 presents the anatomical structure of an MR proposal.

Broadly speaking, the structure of an MR proposal mirrors the MR process, as discussed in popular text books (Kumar, Aaker, and Day, (2009); Malhotra, 2007; Zikmund and Babin, 2007). However, no single framework, no matter how comprehensive, can ever capture all the different ways in which an MR proposal can be put together. Accordingly, Table 2 should be viewed as a prototype, around which the reader can experiment, if necessary. The next several sections will discuss the individual elements in detail.

Table 2 Anatomy of a marketing research proposal.

Title Page
Table of Contents
Executive Summary
Background and Situation Analysis
Research Objectives
Research Design
• Conceptual Framework
• Scope of Research
• Line of Inquiry and Questionnaire Flow
Analytical Plan
Logistics and Execution
• Sample Design and Size
• Quality Control Checks
• Project Schedule
Product and Service Deliverables
Research Investment
Project Execution Team
Why Us?

Cover Page and Table of Contents

As noted above, an MR proposal is both an information and a sales document. Consequently, the cover page should not be viewed as a mere functional necessity. At a minimum, it should be used as a signaling tool and to generate positive momentum from the outset. Perhaps clients are not going to judge a proposal by its cover, but they are definitely going to be unimpressed by the one that does not even try to stand out. Regardless of whether the clients view the proposal as an e-document, or as a printed document, the cover is what they will see first – and first impressions do count.

At a minimum, the cover page should have the following characteristics:

- *Customization*: it should be customized for the client, company, or SBU, inviting the proposal.
- *Specific title*: it should have a clear and tell tale title that offers a window to what the reader can expect from the proposal; example, saying testing service concepts for home-based businesses, is better than merely saying concept test MR proposal.
- *Visual appeal*: we live in an image-rich world; a relevant image, related to the theme of the proposal could greatly enhance the visual appeal of the MR proposal.
- *Date and contact details*: the name and contact details of the primary contact and the date of proposal preparation communicate service friendliness; the reader does not have to shuffle papers to reach the relevant person when they have a question, or need more details.

The table of contents (TOC) is an essential housekeeping chore, not glamorous, but contributing significantly to the overall efficiency and service friendliness of the document. A proposal could be anywhere between 15 and 25 pages long, or longer, depending on the format being used – Word, PowerPoint, or PDF. Searching for relevant content could be very inefficient and time consuming without a TOC page. Imagine a scenario where the client is reading the third of five invited proposals. Assume that the client is intrigued by the proposed research design and wants to revisit the first two proposals to make comparisons. A TOC page will make it easier for the reader to revisit

the relevant sections of the first two proposals without unnecessary shuffling and searching.

Executive Summary

An MR proposal is usually distributed to several key-decision makers within the client company. Not all of them are likely to be motivated to read every page of the proposal. An executive summary (ES) is an indispensable tool for those who want to get a macro understanding of the contents of a proposal without having to read every single page.

An ES should be written assuming that the person reading it will read no other page in the proposal. Well-written executive summaries usually have the following defining characteristics:

- *Completeness*: an ES should be complete, it should cover every single element referred to in the TOC and should read like a stand-alone document.
- *Specificity and concreteness*: there is not enough space in an ES for adjectives and qualifiers like state-of-the art, large, comprehensive, and so on. It is far better to say that 1500 respondents will be contacted, a 20-minute Internet survey is planned, and so on.
- *Brevity*: maximum one page, preferably shorter.

Background and Situation Analysis

MR is a means to an end – it is commissioned to provide information inputs to decision making. Communication between the client commissioning the research and the service provider preparing the proposal is positively served when there is a mutually unequivocal understanding of:

- the decisions most likely to be influenced by the MR project and
- the context within which the decisions are most likely to be made.

The role of this section is to demonstrate this mutually unequivocal understanding of the context and the application.

A common mistake encountered at this stage of the proposal is the tendency to merely reproduce the text contained in the RFP in different words. A mere restatement does not demonstrate an understanding of the client's business situation or decision challenges. To illustrate, one of the biggest casualties of the current financial crisis has been the erosion of consumer trust and confidence in large financial institutions. Now, if you as an MR service provider are hired by clients like Citi or Bank of America to help them regain lost consumer trust, you are more likely to succeed if you display a crisp understanding of the dynamics that led to the erosion of trust. Merely restating that the MR proposal is being written to help clients like Citi and Bank of America regain lost trust of the consumer is not going to inspire much confidence in the reader.

This is a stage-setting phase of the proposal. By definition, therefore, the understanding reflected at this stage should have a clear connection with the content to follow. The content of this section should materially influence research objectives and research design to be presented later. In the absence of such continuity, the holistic-integrated properties of the proposal are likely to be compromised.

It is not uncommon for the MR service providers to conduct some desk research at this stage to reinforce or refine client perspectives reflected in the RFP. It is highly recommended, as it improves the overall thinking about the decision problem and the information to be collected. Continuing with the erosion of consumer trust example, desk research may suggest that experts are recommending a return to neighborhood banking principles. If the client's RFP is focused exclusively on exploring mass media initiatives, you may want to help modify the client's perspective by recommending community-based initiatives as well.

Research Objectives

As stated above, when discussing the mini-case study on obesity, client companies commission MR studies because they want relevant information to make better business decisions. MR service providers write MR proposals to indicate how relevant data and actionable insights will

be provided. Research objectives help ensure that these two perspectives converge. Before the client spends any money, research objectives should help answer the question – *if we execute this study, will we actually get the information we are looking for to help us make better decisions*? Additionally, research objectives also help to audit project performance – did the MR service provider actually deliver on all that the client had contracted for?

As the development of research objectives requires translating managerial information needs to MR tasks, judgment, and skills which come into play in determining the best way to structure the data collection task. What makes it even more challenging is that usually there is no one best way to execute the MR study. A few examples of how business questions were translated to research objectives are given below. It is important to acknowledge that different researchers may have their own points of view on what is the best way to answer the managerial question posed, and may disagree with the recommendations presented below.

- A snack food company wanted to introduce line extensions for its chips and crackers line. The research response to this marketing question was to research salty snack concepts that would meet snacking needs between the hours of 3 p.m. and 6 p.m. (postlunch, predinner).
- Should an insurance company continue to pay a hefty annual amount for sports sponsorship, in this case, an annual tennis tournament? This business question was researched by modeling the return on investment (ROI) on the sports sponsorship expenditure using the customer equity and customer life-time value framework.
- Are all shoppers identical? How do they differ in terms of their planning, their frequency of shopping, their average expenditure the assortment of the items they shop for, and so on? This marketing question was researched by conducting a nationwide shopper segmentation study.

No client or MR service provider has unlimited resources to conduct an MR project. This poses a further challenge to the development of research objectives. The implications of limited resources are obvious – MR professionals need to make trade-offs – in terms of what should be included in the proposed study and what should be left out. This is not always easy, as business and marketing issues are usually influenced by many factors. Trade-offs can be compounded if the research is catering to the information needs of multiple stakeholders and/or if the client is insisting on making the research project carry a heavier data collection load than it is designed for.

Something that is often overlooked in the development and articulation of research objectives is stratifying them as *primary* research objectives and *secondary* research objectives. While not all MR proposals need to consider this stratification, keeping this distinction in mind does help in apportioning emphasis and scarce data collection resources, when the need arises.

For example, a secondary objective in researching patients' treatment practices for migraine could be the actual usage in the past 90 days of alternate therapies, home-made remedies, and over the counter OTC drugs for the treatment of the headache. The secondary objective makes good business sense, because usage of Rx therapies as a class may decline if the usage of alternate, homemade, and OTC therapies increases. This could affect all brands in the category, regardless of how dominant a brand is.

Research Design

In its essence, this section addresses the fundamentals of data collection – how will the MR service provider collect the data, why the provider thinks that it is the best way to do so, and which instruments will be used in the data collection process? Hopefully by now the reader has sensed that the various elements of the proposal are like sequential legs of a relay race, with the preceding section handing off the baton to the succeeding one. The previous leg of the relay, or section, discussed what research objectives will be addressed by the study; this leg of the relay, or section, will discuss which data will address those research objectives and how will those data be collected (*see also* MARKETING RESEARCH PROCESS).

This section should discuss at least three subelements:

- *conceptual framework*: the overriding plan, thinking, framework, and concepts that will guide data collection.
- *scope of research*: establishes the domain and limit of data collection, especially as it applies to geographies and populations.
- *line of inquiry*: the type of data that will be collected, the range of topics that will be covered, and the sequence in which they will be addressed.

Conceptual framework. One of the strongest differentiators between the good research and the unsatisfactory research is the fit between the managerial questions posed by the client and the research solution provided by the service provider. For example, if the question posed by the client is intrinsically longitudinal in nature as in – *how do above the line and below the line marketing expenditures influence brand equity* – then offering a cross-sectional solution will not help, as it will not take into account lagged effects of marketing expenditures. Accordingly, it is essential that the discussion in this section proceeds in a macro-to-micro fashion, where the overall approach is discussed before proceeding to a discussion of specific micro details. Additionally, to the extent possible, details should also be provided about alternatives that were considered and rejected, along with the reasons underlying the preference for the recommended alternative.

An example of macro-to-micro thinking in navigating the research design and conceptual framework follows.

- Address first the overall recommendation, for example, qualitative research, quantitative research, or some combination of the two. If qualitative research is recommended, then give reasons how the qualitative phase will add value and why it is necessary; if not recommended, explain why not necessary or useful.
- Unless, the RFP explicitly asks for an MR proposal for qualitative research, if you are not considering a quantitative research component, you should explain why you feel qualitative research alone will be sufficient in addressing the client's information needs.
- Once the macro aspects of the design have been presented, the discussion should proceed with a discussion of design choices within what is being recommended. For example, if qualitative research is recommended, is the preferred framework in-depth interviews, focus groups, or some form of observational research, like ethnography? As before, pros and cons should be presented in support of your recommendation.
- Similarly, if quantitative research is recommended, will the framework be descriptive research, predictive research, experimental design, or modeling? Again, reasons should be provided in support of what is being recommended.
- For both qualitative and quantitative research designs there should be a discussion on the platform that will be used for data collection, whether in-person, telephone, or Internet.
- Once the above issues have been addressed, the discussion can proceed with further micro details concerning number of interviews, interview length, type of experimental design, types of data to be collected, and so on.

The advantage of this systematic macro-to-micro approach is that at all times the reader has an excellent understanding of the thinking and assumptions guiding the proposal writer's recommendations. The client may not agree, or may have other opinions on what is the best approach, but at least they will understand fully the proposal writer's intentions.

Let us illustrate the above discussion with a mini-case study involving a customer-satisfaction study.

> A cell phone handset manufacturer wishes to conduct an MR study to identify key drivers of customer satisfaction from individual users of cell phones. If the research provider has to deliver on what is expected, the client and/or company A must have a relevant and updated list of satisfaction drivers before measurement can begin. Right away a decision needs to be made. Is the available list of attributes good enough, or

do we need fresh qualitative research to refresh the list? Once that decision has been made, we then need to decide what framework will be used to determine whether or not a driver is key, and which drivers are more important than others. Let us assume that the research provider is leaning toward an importance–performance framework, because in their experience the four-quadrant approach is an effective tool for communicating with marketing decision makers. Once a decision has been made to use the importance–performance framework, additional questions – stated importance or derived importance; first-hand experience of performance or perceived performance – will have to be answered. If the researcher prefers the derived-importance approach, a case should be made for it, including some discussion on why the stated importance approach will not perform as well. Following this, the researcher can discuss issues related to types of attributes, type of data, rating scales, and analytical techniques.

The nested macro-to-micro approach illustrated above has definite merits. It helps make the thinking of the proposal writer visible to the reader. It also ensures that the proposal writer has not made any unwarranted assumptions, such as the list of drivers is current and does not require updating, or omitted any relevant details.

Scope of research. Once issues related to how data will be collected (*see also* WEB SURVEYS) have been articulated, it is time to discuss which geographies data will be collected in and from which respondent populations.

In today's globally interconnected world, it is dangerous to assume that research projects will be conducted in single homogenous geographies. In Europe, Asia, and Australasia, research is routinely conducted in several different markets (read as countries), each with different respondent, consumer, and language characteristics. The increase in minorities (Hispanic, Black, Asian Indian) makes geographic homogeneity a dangerous assumption, even in the United States, even though it is one single country.

This section of the proposal should explicitly address issues related to geography of data collection, languages in which data will be collected, and populations that will be included/excluded from the study. A sampling of issues that should potentially be discussed in this section follow:

- If the study is to be executed in different geographies/countries, will the questionnaire be translated and administered in the country's spoken language? Details should be provided on how this will be achieved.
- Even if the questionnaire is administered in English, as in UK, India, Hong Kong, and Australia, will expressions and rating scale descriptors be adapted for different environments, or will they remain the same?
- Which populations will be included in the study? Returning to the snack food line-extension study, discussed earlier, will respondents be drawn from the general population, from the population of respondents who snack any time of day, or from the population who snack postlunch only?
- If several populations are to be included in the study – say physicians, nurses, and patients in the case of a pharmaceutical drug study – will the same data collection instrument be used for all populations, or will it be adapted? And if adapted, how? Which aspects of data collection will be common and which unique?

Line of inquiry. A line of inquiry indicates to the reader the type of data that will be collected and the sequence in which they will be collected, without getting into the specifics of questionnaire design and question wording. The line of inquiry can play a key role at this stage because it neatly captures the full range of preceding discussion on managerial information needs, research objectives, and research design.

To recount, by the time we get to this stage of the proposal, there should be agreement on which managerial issues we are trying to address, the type of MR data that will best meet the client's information needs, and how best to collect the relevant data. Given this background, the research executive preparing the proposal should be able to specify the types of data to be collected and the sequence in which the researcher thinks they should be collected.

Let us illustrate this with the help of the snack line-extension case study presented earlier. To recap, an experimental design was suggested to help collect the data. Let us assume that the MR executive preparing the proposal disagrees with

the author and prefers instead a survey-based approach. On the basis of the data to be collected by the survey, the executive develops a line of inquiry, presented in Figure 1.

The line of inquiry makes a valuable contribution. It states clearly the data that the proposed study will collect and the sequence in which they will be collected. It also enables the client evaluating the proposal to determine whether any data that should be collected is missing. For example, on reviewing the document, the client may feel that product/brand substitution data are missing; namely, which snack foods and/or brands, if any, will be replaced by the proposed concepts. The client feels that these data should be included in the line of inquiry. Additionally, assume the client is not totally comfortable with each respondent evaluating only two concepts. Why not three? Surely, the respondents can evaluate three concepts without excessive respondent fatigue or without excessively increasing the interview length. Our client and MR service provider preparing the proposal can now have a focused discussion on these concerns, something that would not have been as easy in the absence of a well-documented and articulated line of inquiry.

Analytical Plan

The analytical plan (AP) plays a key role in drawing inferences and insights from the data collected by the MR project. The AP has an intimate link with the research design, as specific analyses require very specific data inputs. To illustrate, for the customer satisfaction case study presented in the previous section, once the set of key drivers is determined and importance–performance data are collected, some analysis will have to be conducted to identify the key drivers of customer satisfaction. If the

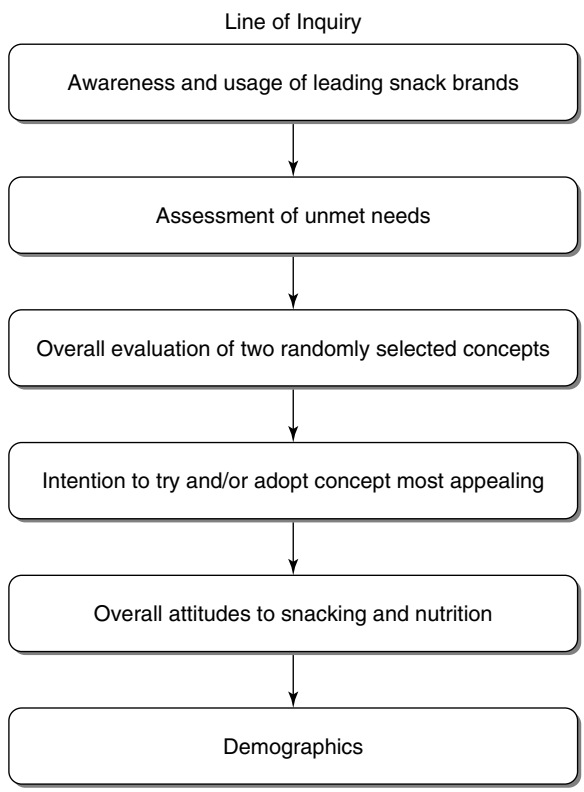


Figure 1 Line of inquiry.

research executive wishes to conduct ordinary least squares, regression to determine which drivers are key, then it is necessary to ensure that interval level data are collected for both sets of independent and dependent variables. If instead of interval data the researcher decides to opt for categorical or check-box data, then the choice of analytical technique will have to be adapted and modified (*see also* UNIVARIATE TECHNIQUES, EXPERIMENTAL DESIGN, ANALYSIS OF VARIANCE AND COVARIANCE, CLUSTER ANALYSIS, EXPLORATORY FACTOR ANALYSIS, MULTIDIMENSIONAL SCALING OF PREFERENCE DATA, DISCRIMINANT ANALYSIS FOR MARKETING RESEARCH APPLICATIONS, MULTIPLE REGRESSION, STRUCTURAL EQUATION MODELING).

Research design and AP are the two sides of the same coin – the former enables the latter. In fact, in some cases the research design virtually handpicks the AP. Longitudinal analysis of advertising effects on sales suggests that the AP will rely heavily on time series or econometric techniques. Similarly, experimental research designs rely very heavily on ANOVA (analysis of variance)-based techniques to quantify the influence of main effects and interactions between the design variables. However, in most cases the researcher does have discretion, in terms of the types of data to collect and consequently, the types of analytical techniques to use. For this reason, the MR executive preparing the proposal is well advised to consider the research design and the AP as a set of interlinked decisions, rather than as two separate decisions.

A few cautionary statements, observations, and myths concerning analysis and APs are presented below:

- Most studies collect a mixture of different types of data – nominal, ordinal, interval, and ratio. The AP should cover proposed analyses for all types of data collected.
- Frequently, analysis is assumed to mean complex or multivariate statistical analysis only. However, analysis refers to all operations performed on the data to make sense of it, including simple summarization and cross-tabulation procedures. All analyses, simple or complex, should be discussed in this section.
- Often, it is assumed that analysis can only be performed on quantitative data. That is incorrect. Qualitative data can and should be analyzed using formal analytical procedures (Miles and Huberman, 1994). Techniques likely to be used to analyze qualitative data should also be discussed in this section.
- Modeling is a common noun for a variety of techniques that depict the structural relationships between the data. If the MR proposal suggests modeling as part of the AP, details should be provided on the type of model that will be developed (linear, nonlinear), data inputs to the model, and statistical techniques that will be used to estimate relationships between the data inputs.
- On most occasions, MR service providers will use off-the-shelf software, like SAS, SPSS, and so on to analyze the data. If the service provider submitting the proposal intends using proprietary software or techniques, details should be provided to make the recommendation as clear as possible (of course, without divulging proprietary details). Black-box approaches, which require the reader to operate on blind trust, may make it difficult for the client to evaluate recommendations made by competing proposals.

Logistics and Execution

This section deals with execution details, such as

- How many respondents will be interviewed?
- How long will each of the interviews last?
- What procedures, if any, will be followed to ensure that the data collected are clean, reliable, and valid?
- How long will the entire study take – from start to finish?

Three topic areas should be covered here:

1. Sample design and sample size.
2. Quality control checks.
3. Project schedule and timing.

The author's preference is to organize the discussion on all execution elements in one section. However, depending on the preference of the executive preparing the MR proposal, it is possible to cover sample design and sample size in the research design section, while discussing populations and subpopulations to be researched.

Sample design and sample size. If the proposal plans to use specialized sampling plans, like random digit dialing (RDD), two-stage sampling, or quota sampling, then an explanation should be provided covering both design and execution details (*see also* SAMPLING TECHNIQUES, NONPROBABILITY SAMPLING; PROBABILITY SAMPLING). As recommended in the research design section, a macro-to-micro approach is very helpful here as well. An example of the recommended macro-to-micro approach in the context of the snack food concept study presented earlier follows:

- *probability* sampling is preferred to *nonprobability* sampling, as greater precision is required to forecast trial and adoption of winning concept;
- within probability sampling, *stratified sampling* is preferred as three subpopulations need to be interviewed – heavy users, light users, and category nonusers;
- stratified sampling *disproportionate to size* is preferred to give a heavier weight to heavy users; if proportionate to size used, we may get more light users and nonusers than needed.

The sample size should be presented for the total study and for each sampling subgroup. Continuing with the above example, the proposal may recommend a total study sample of 1500, split into the three subgroups as follows – 800 heavy users, 450 light users, and 250 nonusers.

In addition to a discussion on sample design and size, *sampling frame* should also be mentioned. How will the respondent's be selected? Will they be drawn from a panel, or will they be selected from lists purchased from list vendors? If the sampling design is to live up to its promise, the sampling frame must play its part; poor sampling frames, no matter how elegant or sophisticated the sampling design, will not yield the quality of data hoped for.

Quality control checks. All MR studies have two types of errors, sampling error and nonsampling error. Sampling errors can be controlled with appropriate research designs; nonsampling errors, on the other hand, can only be controlled by paying meticulous attention to process and detail.

Implementing quality control procedures should be standard practice no matter how skilled or experienced the researcher or the service provider is. A few must-do quality control checks are presented below:

- Often clients will ask service providers to borrow data collection inputs from other sources, for example attribute lists, from previous studies or from published data. There is no harm borrowing, but not without giving some thought to the quality of the data that are being borrowed. Unduly, small sample sizes, poorly worded questions, imbalanced rating scales, and so on should be a cause for concern.
- Debugging a questionnaire, especially if computerized, for ensuring operational integrity.
- Pretesting a questionnaire to ensure comprehension and answerability is good research practice.
- If data are collected in multiple languages, not only should the translation be ratified, but the questionnaire should also be translated back in the language originally written to ensure minimum distortion of meaning and intent.
- Sample quality is a major issue, even when the sample is drawn from panels. MR proposals should address issues related to sample quality with as much specificity as possible. Results and insights provided by the study are only as good as the quality of the sample.
- If the research design calls for human interviewers, details concerning interviewer training and supervision should be presented. The goal of quality control procedures here is to minimize nonsampling error owing to interviewer bias.

- The discussion on quality control should also cover data cleaning and the treatment of missing data. The latter can be especially problematic as it results in unusable sample, and hence a waste of resources.

Neither sampling error nor nonsampling error can be entirely eliminated. But meticulous attention to detail and processes and diligent execution can minimize them.

Project Schedule

All activities from initial consultation with the client to the final deliverables should be presented in the project schedule. If contingencies need to be planned for, they should be discussed separately. An example of a contingency is the need for factoring in the impact of a competitor's product launch on the client's product concepts being evaluated in the study. The data collection may have to be delayed if the competitor's product launch gets delayed. The project schedule for a quantitative study on migraine conducted with neurologists is provided in Figure 2.

Product and Service Deliverables

A research project has both a product and a service component attached to it. An information product in the form of a report containing results and findings will be delivered to the client at the end of the study. However, service

Project schedule

Activities	Duration
Client discussion and finalization of project specifications	One week
Review of prior research	3–5 days
Development of quantitative screener and questionnaire	One week
Client feedback and approval	One week
Programming and debugging questionnaire	One week
Pre-testing and finalization of questionnaire	7–10 days
Data collection and tabulation	Three weeks
Data analysis and presentation development	2–3 weeks
Presentation of results	Two hours
Strategy workshop	Four hours

Figure 2 Project schedule.

is also provided during the course of a project in terms of briefings, keeping the client updated, providing early top line results, and so on. Complete details of both product and service deliverables likely to be provided to the client should be presented in this section.

It is easy to perceive this section to be a simple narrative listing of tables, reports, and other tangibles. What is often overlooked is that the sum of all deliverables, product and service, is the total value the client receives from the MR service provider. The amount that the client will have to pay for this value follows immediately in the next section. By definition, therefore, it should be in the best interests of the MR service provider to provide a complete picture of the price–value proposition represented by the proposal, without resorting, of course, to exaggeration and false claims.

Research Investment

There is a reason why the word investment is preferred to cost. A research project is an investment to augment the decision-making effectiveness of a company. The value of an MR project does not end with the mere act of data collection. Its value is likely to be felt and experienced for some longer period of time.

Before presenting the investment, it is advisable to list once again what the client is getting for their investment. A detailed activity-related break down of the total estimated investment should also be provided. For example, if the MR proposal recommends the development of a web-based simulator in response to an RFP for a conjoint choice-modeling study, the costs associated with the development of the simulator should be presented separately. The accepted practice is to break down the costs by geography, by subpopulation, and by the different value-added activities (data collection, tabulation and analysis, modeling, and so on.) If the MR service provider is willing to provide certain benefits without charging for them, specific mention should be made of it; of both the benefit to be delivered and that there is no charge associated with the delivery. For example, the last activity in the project schedule presented in Figure 2 is a Strategy Workshop. The MR company may be willing to offer this service at no extra charge to the client. This is not uncommon. MR service providers often offer additional value in the interests of building lasting relationships, or to boost their chances of winning the contract.

Project Execution Team – Why us?

As discussed earlier in the article, an MR proposal is both an *information* document and a *sales* document. MR service providers submit proposals because they want to win the contract. Additionally, as discussed in the *Deliverables* section, the client derives both *product* and *service* benefits from an MR project. The company submitting the proposal, its experience and pedigree, the project team supporting the client on the project are all part of the overall value offering.

The purpose of this section is to provide assurance and to sell. The client would like to be assured that their project will be designed and executed by competent, easy to work with professionals and that the overall experience will be enriching and hassle free. The service provider submitting the proposal would like to be rewarded with the project – partly because of the resources spent in developing the proposal and partly because the service provider may genuinely believe it deserves the contract. The project team, the names and faces, of its members and their brief professional biographies provide the assurance. The service provider's credentials, qualifications, its track record, and its experience provide the basis for selling the proposal.

Conclusion

Companies like P&G, Citi, Tesco, and Coke regularly conduct MR projects to provide relevant and timely market information to their decision makers. MR service providers bid for this business by responding to RFPs and submitting proposals. The MR proposal offers competing service providers the opportunity to show case their knowledge and their experience in bidding for the business. In today's complex and uncertain environments, systematic approaches that demonstrate rigorous thinking and skillful application of subject matter expertise to the client's business problem are likely to enjoy a higher

probability of success. It is for this reason that the framework presented in this article emphasizes structure, sequence, and rigor in the development of the MR proposal.

Bibliography

Kumar, V., Aaker, D.A., and Day, G.S. (2009) *Essentials of Marketing Research*, John Wiley & Sons, Inc., New York.

Malhotra, N.K. (2007) *Marketing Research: An Applied Orientation*, Prentice Hall, Upper Saddle River, New Jersey.

Miles, M.B. and Huberman, M.A. (1994) *Qualitative Data Analysis: An Expanded Edition*, Sage Publications, Thousand Oaks.

SAS www.sas.com. Cary, NC.

SPSS www.spss.com. Chicago, IL.

Zikmund, W.G. and Babin, B.J. (2007) *Exploring Marketing Research*, South Western-Cengage, Florence.

research reliability and validity

Rex Yuxing Du

Before we discuss research reliability and validity, it is useful to understand the *true score model*, which provides a framework for understanding the accuracy of any research instrument (Malhotra, 2006). According to the *true score model* $X_O = X_T + X_S + X_R$, where X_T denotes the underlying phenomenon the researcher sets out to investigate, X_O the actual observation made by the researcher regarding X_T. Any discrepancy between the two can be characterized as an error or inaccuracy. The discrepancy between X_T (the truth) and X_O (the observation) can be further divided into X_S (systematic error) and X_R (random error). Systematic error affects the accuracy of a research instrument in a constant way. The impact of random error, on the other hand, is not constant because it is caused by transient factors that affect the research instrument in a haphazard, unpredictable way.

According to the *true score model*, perfect reliability is achieved when $X_R = 0$ (i.e., there is no random error), and perfect validity is achieved when $X_T = X_O$, and $X_S = X_R = 0$ (i.e., there is neither random nor systematic error). In other words, perfect validity implies perfect reliability. It is important to note that, while unreliability implies invalidity, reliability does not in itself imply validity. Put differently, reliability is a necessary, but not sufficient, condition for validity.

Achieving perfect validity, or even merely perfect reliability, is desirable. Unfortunately, in practice, it is impossible to eliminate all possible errors, be it systematic or random. What researchers can do is to devise various methods to assess the nature and extent of these errors, and, if feasible, come up with remedies to minimize them. The rest of the article focuses on the most commonly used approaches in assessing research reliability and validity, especially in the context of measurement (also *see* VALIDITY AND RELIABILITY; Malhotra, 2006, pp. 283–287; Aaker, Kumar, and Day, 2006, pp. 302–304). VALIDITY IN EXPERIMENTATION addresses validity in the context of experimentation.

A *reliable* research instrument should demonstrate consistency, that is, leading to the same observation/conclusion each time it is used with the same subjects under the same condition. Assessing reliability is essentially about gauging the size of the random error. There are several basic operational approaches to do so, including internal consistency, parallel forms, test–retest, and interrater, all of which involve some sort of repetition, because random errors by definition should not reoccur in the same way when the research instrument is applied repeatedly.

Internal-consistency reliability is probably the most frequently used measure of reliability. It refers to how well the multiple items that are designed to capture the same underlying construct yield similar results, which can be estimated in several ways, including average interitem correlation, average item-total correlation, split-half, and Cronbach's alpha.

Average interitem correlation can be derived by first calculating the correlation between each pair of items and then taking the average of all these correlations. For example, for an instrument with six items, one will have 15 different item pairings (and therefore, 15 interitem correlations). A slight variation of average interitem correlation is *average item-total correlation*, which also uses the interitem correlations. The only difference is that, in calculating average item-total correlation, one needs to compute the sum of all the items of the original instrument and use that total as an additional item when calculating the average correlation between the items. The higher the average interitem or item-total correlation, the higher the internal-consistency reliability.

To calculate *split-half reliability*, one randomly divides all items that are intended to measure the same construct into two sets. The entire instrument (i.e., all items from both sets) is administered to a sample of subjects and the total score for each set is then calculated. The split-half reliability is simply the correlation between these two total scores.

Cronbach's alpha (*or alpha coefficient*) is probably the most frequently used estimate of internal consistency and is closely related to split-half reliability. Imagine that one computes one split-half reliability and then randomly divides the items into another two halves and

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

recalculates, and keeps doing this until all possible split-half reliability estimates have been calculated. Cronbach's alpha is mathematically equivalent to the average of all possible split-half reliability estimates; of course, in practice, statistical software packages have more efficient algorithms for calculating it.

To assess *parallel-forms reliability* (or *alternative-forms reliability*), one needs to create two alternative forms. One way to accomplish this is to create a large set of items that address the same construct and then randomly divide them into two sets. Both sets of items are administered to the same sample of subjects. The correlation between the two alternative forms can then be used as an estimate of reliability. One major challenge of this approach is that one has to generate a large number of items that reflect the same underlying construct, which is often not easy in practice. Furthermore, this approach makes the assumption that the randomly divided two sets are parallel or equivalent. Even by chance this sometimes will not be the case.

Parallel-forms reliability is often confused with split-half reliability. With split-half reliability, one intends to have an instrument that is used as a single measurement and only creates random split-halves for the purpose of estimating reliability. By contrast, the parallel-forms approach is typically used in situations where one intends to have two forms that can be used independent of each other and are functionally equivalent, as alternate measures of the same construct. As compared to all the other internal-consistency approaches, the parallel-forms approach shares a major constraint – the researcher must have multiple items designed to measure the same construct.

Test–retest reliability does not require multiple items and can be estimated by applying the same instrument to the same sample on two different occasions. The closer the correlation between the two separate measurements is to one, the more reliable the instrument. The underlying assumption of test–retest reliability is that there is no substantial change in the construct being measured between the two occasions, which is more likely to be true when the amount of time elapsed between the two measurements is short.

Consider the following example. A group of respondents are tested for IQ: each respondent is tested twice, with the two tests being, say, one month apart. One can use the correlation coefficient between the two sets of IQ-scores to assess the test–retest reliability of the instrument. In the ideal case, the before-and-after scores coincide for each respondent, leading to a correlation of one. In reality, this is almost never the case – the scores produced by most respondents would most likely be different from one test to the other. A major limitation of test–retest reliability is the potential "practice effect," that is, respondents "learn" to answer the questions after the first test, which can affect their responses to the same questions in the second test. For example, the IQ-scores may tend to be higher in the retest.

Interrater reliability (or *interobserver reliability*) is one of the best ways to estimate reliability when the research instrument requires observation or human judgment. It refers to the degree to which two raters are being consistent in their observations or judgments using the same instrument. Interrater reliability can be estimated in two ways, depending on whether the measurement is categorical or continuous. If the instrument consists of categories, one can calculate the percent of agreement between the raters. If the instrument is continuous, one can calculate the correlation between the ratings given by the two raters. One might use the interrater approach, especially when one is interested in using a team of raters and wants to make sure they yield consistent results. If one gets a suitably high interrater reliability, one can justify allowing them to work independently. By contrast, one might use the test–retest approach when one has only a single rater.

Each of the above approaches to assessing reliability may lead to different results. In general, the test–retest and interrater reliability estimates tend to be lower in value than the parallel-forms and internal consistency ones because the first two involve measuring at different times or with different raters.

Validity is about whether a research instrument really captures what it is supposed to examine. Major types of validity include face, content, criterion (concurrent and predictive), and construct (convergent, discriminant and nomological).

Face validity (or *consensus validity*) refers to the logical appropriateness of the instrument used.

For example, one might argue that whether a buyer recognizes an advertisement can be used as an indicator of past advertisement exposure. This is logically appropriate (or of face value), and the instrument would therefore appear to have face validity. Face validity is usually established through little more than common sense. (Of course, how well a person recognizes an advertisement might depend on how interested the person was in the advertisement or some factors other than past exposure.)

Content validity refers to the inclusion of all relevant aspects of the underlying construct. For example, a content valid measure of employee satisfaction should cover a full range of experiences and attitudes (e.g., salary, benefits, work-life balance, work hour, nature of work, mobility, bureaucracy, culture, career development opportunities, and so on) that are relevant to the construct of employee satisfaction. Common sense would indicate that a measure of employee satisfaction lacks content validity if it involves only salary and benefits.

In practice, developing a content valid instrument often requires a carefully selected and balanced sample of elements in order to represent the overall construct. This is because of the fact that many constructs are made up of prohibitively long lists of elements and it may be impossible to measure everything. Thus, a selected sample of elements that can represent the majority of all relevant elements is often used to ensure content validity. For example, in an academic examination, the instructor who sets the questions needs to measure whether his/her students have mastered his/her course (overall construct), but he/she cannot test his/her students' knowledge in everything that he/she has taught over the course of a semester – the examination would otherwise last for days! Fortunately, any reasonable instructor would craft an examination containing only a subset of all possible questions to assess his/her students' knowledge. Furthermore, an examination of content validity probably would include questions from each relevant chapter, systematically (but subjectively) selected by the instructor. Taken together, the selected questions should be representative of the course as a whole and do a good job in assessing a student's overall mastery of the course material.

Unfortunately, in marketing research applications, establishing content validity is not as straightforward as setting an examination. The researcher usually needs to first clearly define the construct that he/she intends to study. Then, the researcher needs to develop a comprehensive list of elements relevant to the measurement of the overall construct through a systematic search process (researcher experience, qualitative and secondary researches are often employed to generate the list). Finally, the researcher would select a "best" subset of elements that probably best represents the overall construct. Of course, the whole process is inherently subjective and therefore content validity can, at best, be approximated but not guaranteed.

Criterion validity is established on the basis of certain empirical criteria. In particular, an instrument is said to be of criterion validity if it varies in an expected and significant way with another or another set of variables. In general, there are two forms of criterion validity, concurrent and predictive, depending on the measurement time of the external variables.

Concurrent validity involves relating the measure of interest to external variable(s) taken from the same time period. For example, one may hypothesize that the overall satisfaction of a firm's employees should be tied to the firm's financial performance at the same time period. Concurrent validity would be established if one observes positive correlation between measurements, taken from the same time period, of employee satisfaction and firm financial performance.

Predictive validity refers to how well a measure of interest predicts certain characteristics or behaviors of some related entity *after* the measurement was taken. For example, the mathematics test scores of a class of third graders should be a good predictor of how well each student will perform in fourth grade mathematics. Another example is predictive genetic testing, which uses patients' genetic makeup to predict the likelihood of developing certain diseases. Note that this is distinct from regular medical checkups that test patients' conditions in order to determine if they are currently ill, which establishes the concurrent validity of the medical checkups. Finally, while having predictive validity established seems to

suggest that there is causal relationship between the measured construct and the criterion, it is not always true.

Construct validity refers to whether the researcher is, indeed, measuring the construct he/she sets out to measure. It is often possible for a researcher's instrument to capture a related construct that is actually different from the one that it was intended. For example, a teacher who intends to use an examination to measure his/her students' understanding of the course material may instead measure the students' ability to memorize and regurgitate.

Obviously, in order to establish construct validity the researcher must first have a deep understanding of the underlying construct. The researcher needs to be able to define the construct explicitly and show that the construct's conceptual definition logically connects to its operational definition. Construct validity is often considered the most difficult to assess because of the potential theoretical complexity and the unobservable nature of many of the constructs in marketing research. For example, is share of category requirements a good operational definition of brand loyalty, or is it confounded with some other factors such as needs for variety?

In practice, there are three types of construct validity that are often assessed: convergent, discriminant, and nomological. Establishing all three types of construct validity is difficult and often requires the accumulation of findings from multiple studies measuring similar constructs using similar instruments.

Convergent validity refers to whether a measure follows the same pattern as other measures of the same underlying construct. When a construct is one-dimensional, various measures that are supposed to represent it should be highly correlated with each other. Otherwise, either the measures are poor (invalid) or the construct is multidimensional.

Discriminant validity is the reverse of convergent validity and refers to whether a measure follows a different pattern than other measures that are supposedly *not* reflecting the same underlying construct. A measure of discriminant validity should exhibit low correlation with other measures of constructs that are theoretically unrelated to the construct of interest.

Nomological validity refers to the degree to which a construct behaves as it should within a nomological set (i.e., a collection of constructs that theoretically relate to one another, often represented in a diagrammatic format). Nomological validity is established when the measure of the key construct behaves in the way prescribed by the nomological set.

An excellent metaphor for the *relationship between reliability and validity* is that of shooting a target with a rifle. The center of the target could be thought of as the construct that is being measured. For each subject, taking a measure is like taking a shot at the target. If the instrument (i.e., the rifle) is perfect, one would hit the center of the target every time a subject is measured. Otherwise, one would miss the center.

Figure 1 shows four possible scenarios. The first one, where all the hits missed the center of the target and landed in a small area at the top-right corner, represents an instrument that is reliable but not valid (that is, it is consistent but wrong). Put differently, the instrument is consistently measuring something that is systematically different from the construct that it has been designed to measure. The second, where hits were spread randomly across the target in

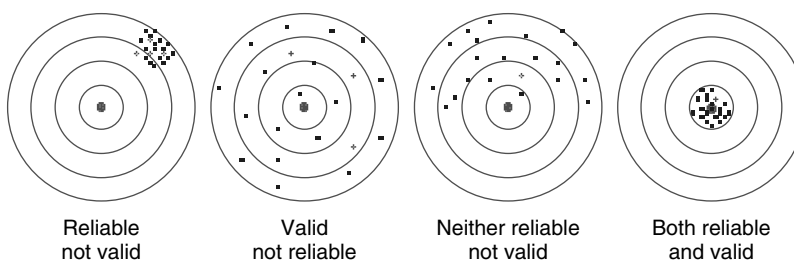


Figure 1 Comparing reliability and validity.

all directions, rarely landing at the center, represents an instrument that is not reliable but may be valid at the aggregate level. In other words, although the instrument gets an erroneous answer from most subjects (i.e., not reflecting the true value of the underlying construct), the errors are not systematically off the center of the target; when aggregated across the sample, the errors balance each other out and as a result, one may get the right (valid) answer for the sample as a whole. That said, the instrument is not valid at the individual level. The third scenario shows a case where the hits were spread across the top half of the target, all missing the center. In this case, the measure is neither reliable nor valid. The last shows a scenario where all the hits landed at the center of the target (or very close to it), representing a measure that is both reliable and valid.

Bibliography

Aaker, D.A., Kumar, V., and Day, G.S. (2006) *Marketing Research*, John Wiley & Sons, Inc., Hoboken.

Malhotra, N.K. (2006) *Marketing Research: An Applied Orientation*, Pearson Prentice Hall, Upper Saddle River.

focus groups and depth interviews

Dennis W. Rook

As long as there have been marketplaces, business people have gathered intelligence about them, in both formal and less structured ways. For centuries, marketplace information has been provided by explorers, scouts, runners, agents, representatives, salesmen, spies, tax gatherers, census takers, and other government functionaries with vital statistics (Levy, 1994). The twentieth century witnessed the steady proliferation of companies, information gathering technologies, and generalized procedures that collectively comprise today's global multibillion-dollar market research industry. Management's ongoing need to obtain information from consumers plays a central role in marketing analysis, strategy, planning, and control. Historically, marketers gathered such data through extensive surveying or polling that generated statistical measures of consumers' demographic characteristics, product and media consumption behavior, purchase intentions, and so on.

This type of information assists managers' strategic and tactical decision making in numerous ways. However, the explanatory power of survey data is limited by its largely descriptive, factual, and somewhat mechanical qualities. The subtler factors that explain why, for example, women like Brand X significantly more than men often lie below the surface and are likely to elude blunt survey instruments with their direct questions, fixed-format responses, and few degrees of freedom. When grumbling about survey research was gaining momentum in the 1930s, psychological theories, insights and methods had begun to filter into marketing thought and quickly expanded and enriched marketing's intellectual and substantive domain. A watershed in this period was an article by the renowned psychologist Paul Lazarsfeld (1935): "The Art of Asking Why in Marketing Research." Some scholars view this article as the seminal force in opening the door to behavioral science thinking and methods in marketing research (Levy, 1994). No longer wedded so strongly to quantitative survey methodology, various types of what came to be known as *qualitative research* emerged onto the marketing scene.

Within the qualitative research arena (*see* EXPLORATORY RESEARCH; MOTIVATION RESEARCH; OBSERVATION METHODS), individual marketing researchers designed their studies around talking with and interviewing consumers either on an individual basis or in group discussions. Today various technologies facilitate the collection of information about consumers' shopping activities (scanner data, *see* ELECTRONIC SCANNER SERVICES) and media consumption (Nielsen data, *see* PURCHASE, SCANNER, AND MEDIA PANELS) electronically and almost invisibly, with minimal consumer involvement. These technological advances, however, hardly eliminate marketers' needs to talk with consumers live and in person. One-on-one interviews are most commonly called *individual depth interviews* (*IDIs*), and group discussions are described as focus groups. Although, they share common origins, IDIs and focus groups are quite different in significant ways and have various strengths and weaknesses.

Individual Depth Interviews

The following material presents a brief history of IDIs and identifies their distinctive qualities and uses. Then the discussion considers the principles that guide the design, structure and gathering of data from individual consumers and other marketplace intermediaries such as purchasing agents, wholesale and retail personnel, and others.

Origins and evolution. IDIs migrated into the marketing-research toolkit from the fields of clinical psychology, qualitative sociology, and cultural anthropology. Arguably the single strongest influence came initially from clinical psychology, whose nondirective individual interviewing and imaginative projective techniques were easily and successfully adapted for commercial purposes. In the early days, roughly from 1935 to 1960, market research uses of IDIs were frequently guided by Freudian and neo-Freudian thought. Perhaps more than any psychologist ever, Sigmund Freud's compelling and provocative theories achieved

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

deep penetration into the popular culture during this period, and filtered down into cocktail party chatter about the symbolic meaning of cigars and convertibles, among other things. These developments represent an historic intellectual and managerial shift from an emphasis on counting and measuring things toward the then exciting and uncharted research territory involving consumers' motivations and meaning systems. In fact, most scholars refer to this period as the "Motivation Research" era (*see* MODELS FOR CATEGORICAL DEPENDENT VARIABLES). In the 1950s, a flurry of academic articles, trade publications and books discussed the concepts, tools, and findings of motivation research, which overwhelmingly relied on IDIs (Martineau, 1957; Newman, 1957; Smith, 1954).

Depth interviews' nature and uses. The primary purpose of a depth interview is to achieve understanding that goes beyond the obvious and surface-level to reveal consumers' basic values, predispositions, needs, feelings, and motivations. For example, when a consumer is asked to explain his reasons for purchasing his current car, a frequently obtained initial response is: "I liked its look and I got a good deal on the price." This rationalized explanation actually provides very little information, as few consumers would be likely to report wanting to buy ugly and overpriced cars. Additional probing might, hypothetically, yield the finding that this particular 46-year-old male consumer purchased a convertible sports car to signal his financial success and sexual availability. Clearly, even the most self-aware consumer is unlikely to offer such an intimate explanation initially, so it must be teased out during a lengthy interview process that relies on both direct and more subtle approaches to the question(s).

Probably the most renowned among the early migrants from clinical psychology to marketing research was Paul Lazarsfeld's student Ernest Dichter, who fled Nazi Europe in the 1930s and set up business in New York, not as a clinical therapist but as a corporate consultant. He achieved early success and enjoyed a long consulting career with America's top businesses, and his company also provided a training ground for young qualitative researchers, many of whom went on to have their own successful, independent careers.

A signature of Dichter's research was its analytic reliance on the most basic human dialectics that cut across issues and tensions involving aspects of gender, social status, age, and (especially) sexuality, along with the attendant interpretation of symbolic meanings that swirl around consumption, products, and brands. Anchoring market research in such "deep" topics and concerns is one identifying aspect of IDIs. Dichter's numerous studies collectively provide a prototype of what "depth" research is. His findings and interpretations were almost always fresh, original, and thought provoking; sometimes they were quite sensational and controversial, and filtered into the media and public discourse. His *Handbook of Consumer Motivations* (Dichter, 1964) is a fascinating compilation of many dozens of studies across the marketplace landscape. Excerpts from two client studies are included to illustrate his analytic approach:

> Study 1: Baking does not play the same role in the life of the modern woman that it did even as recently as one or two generations ago . . . The modern woman no longer has to bake of necessity . . . however, she continues to take an interest in baking because of the emotional satisfactions which she derives from this activity. Baking is an expression of femininity (that) evokes pleasant, nostalgic memories . . . (and) provides a feeling of love and security . . . The most fertile moment occurs when the woman pulls the finished cake, bread loaf, or other baked product out of the oven. In a sense it is like giving birth. How did it turn out? Did it fulfill expectations? (p. 21)
>
> Study 2: There is probably a major difference between . . . men who use electric shavers and those who use razor blades (that) is anchored in the basic psychological significance of shaving. Each time a man shaves he is, on the one hand, robbing himself of a masculinity symbol, on the other hand reasserting it . . . The shaver who uses razor blades is a person who enjoys the ritual that goes with it . . . Razor blade users who have tried electric shavers often rationalize that they do not get a close enough shave. While this may be true, it is also very likely that a deeper reason for not favoring a dry shave is that it eliminates much of the elaborate and somewhat masochistic ritual of wet shaving. (p. 184)

Some people might find the comparison of baking to giving birth utterly preposterous, yet the Pillsbury Company seemed to appreciate the connection, and for many years their advertising reminded American housewives that "Nothin' says lovin' like somethin' from the oven, and Pillsbury says it best." Similarly, Gillette seems well aware of men's ambivalence about the ritual of shaving. Their long-term sponsorship in the 1950s and 1960s of TV's Friday Night Fights positions their razors in an ultramasculine venue, while their more recent slogan, "Gillette: the Best a Man Can Get," appeals to contemporary metrosexual consumer interests and motivations.

Technique: Structure and Procedures

Several distinctive structural and procedural qualities of an IDI facilitate achieving the enhanced level of understanding that the technique's pioneers intended. One of the central elements of depth interviews is simply assumed and rarely discussed: the *one-on-one* nature of the questions and conversations. Individuals today rarely receive the undivided attention of another for any significant length of time except when they pay for it in the professional services domain. This feature makes depth interviews experientially unique and tends to induce feelings of intimacy and privacy, which further encourages greater self-disclosure than one would obtain in either survey or focus-group research. Depth interviews are also generally *lengthy*, lasting between 1 and 2 hours, a characteristic embedded in the title of Grant McCracken's delightful book, *The Long Interview* (McCracken, 1988). Surveys and even focus groups often seem like interrogations rather than interviews, with rapid-fire direct questions posed to consumers about a myriad of topics, with the expectation that responses will be made quickly and without much elaboration. Depth interviews are, or at least should be, in much less of a hurry to extract unambiguous answers to specific questions.

Another characteristic of depth interviews is their reliance on *open-ended questions* that offer greater degrees of response freedom, which encourages more thinking and reflection on the consumer's part. For example, asking bluntly "Do you like Nike?" would tend to elicit a simple yes or no response, and require additional probing about the underlying reasoning. A depth interview would open up and broaden, essentially, the same question by asking, "Tell me how you feel about Nike." Open-ended questions not only capture a broader range of responses, but they also make for interviews that are more *conversational* in nature. This does not, however, imply that the dialogue is the same as that between friends and family, with the two parties going back and forth. Reflecting their clinical psychological origins, the depth interviewer asks a question and lets the respondent run with it without much intervention, often merely nodding or making brief verbal comments to encourage continued explanation beyond the initial response. When an individual seems to have exhausted the topic, veered off into tangential territory, or simply run out of steam, the interviewer moves on to the next topic or question.

Being conversational in nature does not imply that depth interviews are without clear and purposeful structure. Like focus groups, they tend to *flow from general to increasingly specific* questions. To illustrate, in a research project sponsored by a soy sauce manufacturer some years ago the interview discussion guide was designed to cover eight increasingly narrow thematic topics, with several questions under each. The interview began with three questions about the respondents' (1) overall lifestyles and four questions about (2) their cooking and meal preparation patterns and activities. This was followed with four questions about (3) "ethnic" food trial and likes, three questions about (4) "Oriental" food, and then nine questions about consumers' (5) understanding of, first experiences with, and current uses of soy sauce. Then four questions probed their (6) perceptions of different soy sauce brands. The interview concluded with the collection of consumers' (7) evaluations of concepts statements for new soy sauce products and of (8) different advertising storyboards. A condensed sample of the discussion guide is included in the Table 1. Despite the study's obvious commercial purposes, the nature and flow of the topics and questions provide the basis for a relatively naturalistic consumer conversation. Although, some managers will likely obsess

Table 1 Soy sauce depth interview discussion guide.

Overall Lifestyle
1. First, please tell me a bit about yourself . . . your work . . . family.
2. What are your special interests . . . leisure activities?
3. How do you think your lifestyle is different from your friends?

Cooking and Meal Preparation
4. Do you like to cook . . . compared to your friends and relatives?
5. What kinds of dishes do you make on a regular basis . . . why these?
6. What kinds of dishes do you make for special occasions?
7. What kinds of meals do you enjoy cooking the most . . . why?

Ethnic Food
8. What kinds of ethnic food you especially like . . . what is its appeal?
9. In your own cooking, do you experiment with ethnic or foreign foods?
10. Tell me about some ethnic dishes you have made recently. How did they turn out?
11. Have your tastes changed in the kinds of ethnic foods you like . . . why?

Oriental Food
12. What are your feelings about Oriental food . . . likes . . . dislikes?
13. What are your favorite Oriental dishes . . . least favorite?
14. Are you making more or fewer Oriental dishes than you used to . . . why?

Soy Sauce: Basic Understanding and First Experience
15. You said you use soy sauce. Where does it come from . . . where is it used?
16. Soy sauce is made in China, Japan and the United States. Does it matter to you where it comes from?
17. What was your first experience with soy sauce . . . where . . . when? How did you like it?

Soy Sauce: Purchase and Usage
18. How often do you buy soy sauce? Is it a planned or impulse purchase?
19. Tell me about the dishes you make that have soy sauce as an ingredient.
20. Is soy sauce more a part of your everyday cooking or is it more for special occasion meals?
21. What are some non-Oriental dishes in which you put soy sauce?
22. Do you use soy sauce more or less often than you used to . . . why?
23. Do you also use soy sauce as a condiment on the dining table?

Soy Sauce: Brands
24. What soy sauce brand do you usually use . . . why? What makes it better than other brands?
25. What other brands of soy sauce are you aware of? What do you like or dislike about each of these?
26. If the store was out of your favorite brand, what would you buy instead?
27. Here is a picture of two women grocery shopping. One woman prefers La Choy soy sauce and the other prefers Kikkoman. Tell me a story about what each one says about her favorite brand . . . and what are the two women like?

Soy Sauce: Product Concepts and Advertising Evaluations
28. New soy sauce product concept evaluations (not available)
29. Advertising copy testing

about the product concept and advertising "test" results (questions 28 and 29), more enlightened ones will appreciate and use the experiential depth that lies behind consumers' evaluative responses.

Although they depend heavily on open-ended questions, depth interviews also rely frequently on the unique data elicitation qualities of *projective methods* to tease out answers to questions that more direct approaches often fail to provide (*see* PROJECTIVE TECHNIQUES). Projective methods, also with clinical psychological origins, are a diverse family of techniques (story telling, sentence completion, picture drawing, word

association, dream exercises, symbol matching, among others) that are particularly useful in situations in which consumers for various reasons are likely to be unable or unwilling to articulate the information that researchers need (Rook, 2006). Question 27 in the Table 1 is an abbreviated example of the thematic story telling technique. The creative, constructive, and indirect nature of projective techniques, along with their high degrees of response freedom, permit access to deeper levels of the human psyche where consumers' emotions, visual imagery, impulses, fantasy, and conflicts reside. Finally, depth interviews generate a *large amount of narrative and other expressive data* compared to most market research protocols. The large amount of data offers rich perspectives for interpreting consumers' stories and their underlying meanings. When sample sizes are large the amount of data can seem overwhelming, so researchers have developed software programs (NUD*IST, Ethnograph) for coding and sorting responses, although it is not yet clear that applying quantitative content analytic tools to qualitative data is particularly appropriate or incrementally productive. For a broader discussion of the analysis of qualitative data.

Focus Groups

Although, Ernest Dichter preferred individual over group interviews, his consulting company trained a large number of the first generation of focus-group researchers. Many other influential scholars and practitioners were involved, whose involvement in the evolution of focus-group interviewing is covered below. Following this, the discussion considers the nature and uses of focus-group research as well as the main procedural details (*see also* EXPLORATORY RESEARCH).

Origins and evolution. Originally called *"focused" interviews* (Merton and Kendall, 1946), this technique came into vogue after World War II and has been a part of the social scientist's tool kit ever since. Researchers in numerous behavioral science disciplines have relied on focus groups as a source of primary data. The fields that have embraced, at various points in time, research with groups include education, sociology, communications, the health sciences, organization behavior, program evaluation, psychotherapy, social psychology, gerontology, political science, policy research, and marketing. At the beginning, the theoretical underpinnings of focus groups emerged from what pioneer Alfred Goldman describes as a "rich stew of sociopsychological and psychotherapeutic traditions and techniques" (Goldman and McDonald, 1987, p. 3).

The successful uses of focus groups in evaluating World War II morale and training films did not go unnoticed by the marketing-research community. The procedures developed by Lazarsfeld and Merton were imported directly into CBS' research of pilot radio and television programs, and are still used today. Although, it is likely that some innovative business studies used focus groups as early as the 1930s, their popularity grew dramatically from the 1950s onward (Leonhard, 1967; Smith, 1954). Marketing researchers quickly discovered the versatility of the focus groups in addressing numerous concerns related to designing products and services; obtaining consumers' perceptions of prices, brands, and retail environments; and their reactions to advertising and other marketing stimuli. Also, there is some speculation that, in the days before the rise of commercial interviewing facilities, interviewers grew weary of lugging around heavy reel-to-reel tape recorders for capturing their in-home conversations with consumers, and responded positively to the idea of sitting down to interview a group of housewives in someone's family room.

In comparison to statistical research, focus groups are more user friendly, and can be fielded and analyzed relatively quickly. They also provide an office get-away that is often social and entertaining, as well as insightful. Just like IDIs, the emergence of focus-group theory and practice is linked to the motivation research era of the 1950s (Smith, 1954), also with a strong intellectual grounding in Freudian and neo-Freudian thought. Groups conducted in this tradition tend to share the exploratory, interactive, playful, and confrontational qualities of clinical psychological groups. By contrast, focus groups rooted in social psychological thinking tend to be more evaluative in purpose, direct

in questioning, and lower in respondent interaction. Such focus groups are often heavily involved in gathering consumers' reactions to product concepts, marketing communications, and competitive brands.

The use of focus groups expanded steadily since they were first introduced, and the pace increased over the past quarter century. Perhaps the most compelling quality of focus groups is their delivery of live flesh and blood consumers for observation by marketing managers who are typically hidden behind a one-way mirror, unseen by the participants. This sharply differentiates focus groups from IDIs, which managers typically have little interest in viewing, since this would be far more time consuming and lacking in the entertainment value of group discussions and interactions. Today focus groups account for as much as 80% of the $1.1 billion spent domestically on all varieties of qualitative research (author's estimate). Equally significant, focus groups' share of mind is extremely high, as many managers virtually equate "focus groups" with "qualitative research," having little awareness of other qualitative techniques.

Focus-group nature and uses. A focus group involves a group discussion of a topic that is the "focus" of the conversation. The contemporary focus-group interview generally involves 8–12 individuals who discuss a particular topic under the direction of a professional moderator, who promotes interaction and assures that the discussion remains on the topic of interest. A typical focus-group session will last from 1 1/2 to 2 1/2 hours. The most common purpose of a focus-group interview is to stimulate an in-depth exploration of a topic about which relatively little is known. Focus-group research is uniquely suited for quickly identifying qualitative similarities and differences among people. Focus groups also provide an efficient means for determining the natural everyday language people use when thinking and talking about specific issues and objects, and for suggesting a range of hypotheses about the topic of interest. Focus groups may be useful at virtually any point in a research program, but they are particularly useful for exploratory research of uncharted territory (*see* EXPLORATORY RESEARCH). As a result, focus groups tend to be used early in research projects, and are often followed by other types of research that provide more quantifiable data from larger respondent samples.

Focus groups have also been proven useful following analyses of large-scale, quantitative surveys. In these instances, focus-group research facilitates more nuanced interpretations of quantitative results and adds depth to the responses obtained in the more structured survey. Focus groups also have a place as a confirmatory method that may be used for testing hypotheses. This application may arise when the researcher has strong reasons to believe a hypothesis is correct, and where disconfirmation by even a small group would tend to result in rejection of the hypothesis.

Focus groups can produce quantitative data, but this is at odds with their nature and primary purpose, which is the collection and analysis of qualitative data. Like IDIs, focus groups, when properly designed and conducted, generate a rich body of data expressed in the respondents' own words. In focus groups participants can qualify their responses or identity important contingencies associated with their answers. Thus, responses have a certain ecological validity not found in traditional survey research. On the other hand, the data provided by focus groups may be idiosyncratic and unique to the group.

Although, focus groups can be conducted in a variety of sites, ranging from homes to offices, they are typically held in commercial facilities designed especially for focus group interviewing. Such facilities provide one-way mirrors and viewing rooms where observers may unobtrusively observe an interview in progress. Focus-group facilities may also include equipment for audio- or videotaping interviews and perhaps even small receivers for moderators to wear in their ears, so that observers may speak to them and thus provide input into interviews. In an age of on-line communication and video-conferencing focus-group facilities also tend to be equipped for "virtual" groups where the members may be broadly dispersed geographically and communicate through electronic media. Focus-group facilities tend to be situated either in locations that are easy to get to, such as just off a major commuter traffic artery, or in places like shopping malls, where people tend naturally to gather.

Technique: structure and procedures. The first characteristic is that of *focus*, which was first introduced by pioneers Merton and Kendall (1946). They explain that the basic purpose of the "focussed" interview (their preferred spelling) was to gather qualitative data from individuals who have experienced some "particular concrete situation," which serves as the focus of the interview (p. 541). Alfred Goldman is prominent among the second generation of researchers who transitioned from clinical to marketing research uses of focus groups, and his article, "The Group Depth Interview" (Goldman, 1962) is widely considered a definitive classic that identifies three additional theoretical pillars of focus-group research. A second signature element is the objective to better understand the *group dynamics* that affect individuals' perceptions, information processing, and decision making about the focal research topic. This is the main logic for conducting the research in a group rather than an individual setting, to allow observations of how and why individuals accept or reject others' ideas. Also, stimulating interactions among group participants are hypothesized to generate more information than individual interviews would provide, although there is little empirical support for this position.

A theoretical quality they share with IDIs both in purpose and name is the elicitation of *in-depth data*. A main thread that connects the diverse family of today's focus-group users and providers is a belief that live encounters with groups of people will yield incremental answers to questions that go beyond the level of surface explanation. Clinical psychology is rich with qualitative research tools and techniques such as projective methods and group involvement techniques that elicit the emotions, associations, and motivations that influence particular behaviors. Focus groups in the health sciences often address emotional, even life-and-death issues. Market researchers have similar, although generally less serious concerns about identifying the underlying behavioral factors that account for consumers' attitudes, preferences, and motivations.

The history of focus-group theory and practice is part of the larger history of qualitative research in the behavioral sciences. In comparison to most quantitative survey research, qualitative research is a contact sport, requiring some degree of immersion into individuals' lives. This and its emphasis on meaning rather than measurement have contributed to its characterization as *humanistic* research. This is not meant to ennoble qualitative research; rather it simply points to a general orientation that includes empathy, openness, active listening, observations, and various types of interactions with research participants, which is markedly and experientially different from the prophylactic rating scales that survey researchers deploy.

Focus groups in marketing today tend to perform less well on the humanistic criterion. This is partly because of the tendency to field research more for evaluative than developmental or exploratory purposes, reflecting marketers' voracious appetites for obtaining consumers' critical reactions to new product concepts, advertising copy, and competitive brands. This is understandable and necessary, but such interests might – ironically – be better served through survey research rather than focus-group studies. Also, groups that are dedicated to evaluative polling tend to exhibit characteristics of a business meeting in which moderators misguidedly seek to achieve group consensus. Another factor that is likely to diminish a group's humanistic qualities is the erroneous, but widespread belief that the moderator should ask every question that appears in the discussion guide, which often "destroys the elements of freedom and variability within the interview" (McCracken, 1988, p. 25).

In their early days in marketing research, the prototypic focus group was characterized by a relatively small number of loosely structured questions that center around a focal topic or stimulus, and encourage extensive discussion and probing. Structurally, focus groups and IDIs were quite similar. Overtime, focus groups have drifted away from their original emphasis on achieving in-depth consumer insights. Robert Merton's own impression of focus groups in marketing concludes that they are "being mercilessly misused" (Merton, 1987, p. 557). Two factors have contributed to the decline in focus-group depth. First, focus-group discussion guides tend to include too many questions, which often makes the

experience more like a within-group survey than an interactive discussion. Rook (2003) quantifies the interaction between the number of questions in the discussion guide and the focus-group length. It is common today to have 30 or more questions, which can reduce the response time per respondent to 13 seconds or less. In these circumstances, the moderator is likely to feel hurried and unable to probe interesting or unclear responses, all of which militate against achieving in-depth information.

A second problem emerges from the tendency to use exclusively direct questions and verbal responses to them. This is not only inconsistent with the historical nature of focus groups it also defies current neuroscientific understandings about the workings of the human mind. These findings conclude that the vast majority of human thought is visual, metaphorical, and emotional, and resides deeply in neurological substrata. Access to these mental zones typically requires more subtle, indirect approaches to asking questions; and it suggests using nonverbal techniques that involve visualizations or role playing. Marketing professionals increasingly recognize the problem and the criticism of the focus-group research been growing for a decade. Some companies (Yahoo, America Online) have abandoned them almost entirely in favor of alternatives such as ethnography, which facilitates greater immersion into consumers' lives (*see* ETHNOGRAPIC RESEARCH). The controversy peaked recently when best selling author Malcom Gladwell, in his 2005 keynote address at the American Association of Advertising Agencies' summer conference recommended that focus-groups research should be entirely abandoned. This caused quite a stir, since advertisement agencies are extremely heavy users of focus groups. The ebb and flow of focus-group research across and within various disciplinary fields, and the attendant intellectual elements of thesis, antithesis, and synthesis make focus groups an interesting and dynamic arena that continues to merit further consideration.

One problem that might not seem apparent in the literature, but looms large in the field is what Grant McCracken observes as an intellectual climate that reflects "substantially more concern with practice than theory" (McCracken, 1988, p. 15). This is particularly the case in marketing research, where dozens of articles and books tend to emphasize the dos-and-don'ts surrounding the myriad of executional details involving recruiting participants, preparing discussion guides, selecting moderators, blocking time slots, inviting observers, ordering food, data analysis, and report preparation. As Rook (2003) suggests, the stage management aspects of focus groups often preoccupy researchers to the point where more basic issues are barely considered. In practice, researchers rarely step back to ask why they want to conduct research with groups rather than individuals, and why in a mirrored room instead of a natural setting? Greater consideration of these critical research design questions might invigorate the intellectual environment and reduce the perceptions of focus groups as research junk food.

Conclusion

Over 60 years ago, what we today call IDIs migrated into the marketing field and were quickly added to the marketing-research toolkit. The rapidity of their adoption was influenced by their intellectual partnership with the field of motivation research and its own growing popularity that was fueled by a steady stream of original, analytic insights that marketing managers found useful and actionable. IDIs were also promoted by the extraordinary larger than life research pioneers who sold them to innovative and courageous clients, and published their work in both trade and scholarly journals, including the *Journal of Marketing* and *Harvard Business Review*. Eventually, the neo-Freudian thinking that motivated much individual depth interviewing came under attack from scholars who argued in favor of cognitive and social psychological approaches and condemned the "excesses" of psychodynamic interpretations.

By the 1970s, focus-group research began to take off rapidly as IDIs ebbed. Focus groups had many selling points; the efficiency of a 2- to 4-hour data collection, the quick research turn-around this permits, the user-friendliness of the verbatim data, the voyeuristic aspects of watching people behind a mirror, the ecological validity of listening to live consumers, the entertainment value focus groups commonly provide, and the highly social nature of the event. These

factors and others contributed not merely to the popularity of focus-group research, but to the emergence of a focus-group industry comprised of over 1500 brick-and-mortar facilities for conducting the groups and many thousands of professionals who work in focus-group recruitment, recording and servicing (e.g., food and beverage; transportation, and hospitality). Compared to IDIs much focus-group research today is far afield from its theoretical and methodological roots, which may be inevitable with such mass appeal and usage. The criticism has been growing but it has also been helpful in providing suggestions for improving the quality of the focus-group research by using it in the ways it was intended, and continuing to innovate different focus-group techniques, formats, and locations. The good news is that with the explosive growth of ethnographic research and the renewed interest in projective methods, the next generation of marketing-research professionals will have a far more diverse toolkit to address marketing's far flung concerns and questions.

Bibliography

Dichter, E. (1964) *Handbook of Consumer Motivations*, McGraw-Hill, New York.

Goldman, A.E. (1962) The group depth interview. *Journal of Marketing*, **26**, 61–68.

Goldman, A.E. and McDonald, S.S. (1987) *The Group Depth Interview*, Prentice-Hall, Inc., Englewood Cliffs.

Lazarsfeld, P. (1935) The art of asking why in market research. *The National Marketing Review*, **1**(1), 26–38.

Leonhard, D. (1967) *The Human Equation in Marketing Research*, American Management Association, Inc., New York.

Levy, S.J. (1994) What is qualitative research? in *Marketing Manager's Handbook* (eds S.J. Levy, G. Frerichs, and H. Gordon), The Dartnell Corporation, Chicago, pp. 271–286.

Martineau, P. (1957) *Motivation in Advertising*, McGraw-Hill, New York.

McCracken, G. (1988) *The Long Interview*, Sage Publications, Newbury Park.

Merton, R.K. (1987) Focused interviews and focus groups: continuities and discontinuities. *Public Opinion Quarterly*, **51**, 550–566.

Merton, R.K. and Kendall, P.L. (1946) *The Focussed Interview*, Free Press, New York.

Newman, J. (1957) *Motivation Research and Marketing Management*, Harvard University, Cambridge.

Reynolds, T.J. and Gutman, J. (1988) Laddering theory, method, analysis, and interpretation. *Journal of Advertising Research*, **28**, 11–34.

Rook, D.W. (2003) Out-of-focus groups. *Marketing Research*, summer, 10–15.

Rook, D.W. (2006) Let's pretend: projective methods reconsidered, in *Handbook of Qualitative Research Methods* (eds R. Belk), Elgar Publishing, Northampton, pp. 143–155.

Smith, G.H. (1954) *Motivation Research in Advertising and Marketing*, McGraw-Hill Book Company, New York.

Stewart, D.W., Shamdasani, P.N., and Rook, D.W. (2007) *Focus Groups: Theory and Practice*, Sage Publications, Thousand Oaks.

Wansink, B. (2000) New techniques to generate key marketing insights. *Marketing Research*, **12** (2), 28–36.

itemized rating scales (Likert, semantic differential, and Stapel)

Gary J. Russell

Introduction

One of the key challenges in marketing research is developing effective measures to describe consumer decision making. Researchers are interested in variables such as perceptions of product attributes, the relative importance of attributes in making choices, and preferences for different products. Measuring variables of this sort enhances managerial understanding of perceived product substitutability and the positioning of products relative to consumer segments. By tracking changes in consumer variables over time, managers can gain a better understanding of the determinants of important strategic constructs such as product awareness, consideration sets, brand equity, and market structure.

From a theoretical perspective, consumer variables that measure psychological reactions to the marketing environment are known as *attitudes*. Briefly stated, an attitude is a "psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor" (Eagly and Chaiken, 1993). Attitude measurement is very challenging for the simple reason that attitudes are not directly observable. For example, a researcher could infer that advertising messages strengthen brand positioning by measuring how consumer perceptions of product substitutability change in response to a new advertising campaign. Generally, the only way that a researcher can gain an understanding of these perceptions is by asking a representative group of consumers a set of structured questions. The success of the research will obviously depend upon which questions are asked, how these questions are constructed, and how the answers are interpreted.

Although many types of questions are possible, the focus of this article is a type of question known as an *itemized rating scale*. The phrase *rating scale* refers to a question in which the consumer is asked to indicate the strength of his or her attitude with respect to some entity (such as a product). The word *itemized* refers to the fact that the consumer is allowed only to pick one of K different responses. These responses are arranged on the questionnaire in a way that suggests an ordinal set of relationships. If we think of an attitude as a continuous variable, then an itemized rating scale is an attempt to divide the attitude continuum into K ordered intervals (each corresponding to one of the K possible responses of the rating scale). In response to a question, the consumer reports the interval in which his or her attitude lies.

We first discuss a typology of measurement that identifies the universe of possible questions in survey research. We then review different types of itemized rating scales, and discuss practical issues of scale construction and interpretation. We conclude with a brief overview of modern methods of scale analysis that infer a continuous measure of attitude from the rating scale responses.

Levels of Measurement

Psychometricians have long known that not all data collection procedures yield information with the same statistical properties. Following Stevens (1946), four different levels of measurement can be defined: *nominal*, *ordinal*, *interval*, and *ratio*. All possible questions that might be posed on a questionnaire can be classified into one of these groups (Table 1).

Stevens (1946) defines measurement as the assignment of numbers to objects (*see* PRIMARY SCALES OF MEASUREMENT). The weakest of these assignment rules is a nominal scale. A scale with nominal properties simply classifies entities into groups. A common example is household identification numbers in a panel dataset. The numbers serve to distinguish households, but otherwise have no meaning. On a survey questionnaire, a multiple choice response (such as marital status or gender) is considered a nominal scale. From a statistical point of view, nominal scales are usually called *categorical variables*. Relationships among nominal scales can be explored using statistical procedures designed for contingency tables (Bishop, Fienberg, and Holland, 1977).

A stronger level of measurement is the ordinal scale. As suggested by the name, ordinal scales allow the researcher to rank order

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Levels of measurement.

Type of Scale	*Distinguishing Characteristic*	*Allowable Transformations*	*Allowable Statistical Analysis*
Nominal	Naming outcomes	Permutation of labels	Frequency distribution log-linear models contingency table statistics
Ordinal	Ordering outcomes	Monotonic	Percentiles ordinal regression nonparametric statistics
Interval	Comparison of differences	$Y = a + bX$, where $b > 0$	Means and variances general linear model
Ratio	Comparison of ratios	$Y = bX$, where $b > 0$	Coefficient of variation percentage changes

The four levels of measurement are ordered from nominal (low) to ratio (high). Higher levels of measurement possess all the properties of lower levels of measurement.

entities with respect to some characteristic. For example, marketing research surveys often ask respondents to rank order N products in terms of preference, using the scheme of 1 (high) to N (low). From this information, we can always infer which of two products is preferred more. However, we cannot tell how much more. Differences and ratios on an ordinal scale have no meaning. Specialized analysis procedures, known as *nonparametric statistics*, have been developed to analyze relationships among ordinal variables. Because itemized rating scales have ordinal properties, nonparametric methods can be quite useful in applied work.

Interval level measurements allow the researcher to compare differences between two points, but do not allow ratio comparisons. For example, consider centigrade and Fahrenheit temperature measurements, both of which are regarded as interval scales. Both measure heat, and are related by the linear expression $C = (F - 32)[5/9]$. Note that, the definition of one degree is different across the two scales, and the freezing points of water are also different. This arbitrariness in fixing the unit and the zero point of an interval scale makes ratio comparisons meaningless. Virtually all common statistical procedures (such as regression analysis) assume that the dependent variable has all the properties of an interval scale.

The highest level of measurement is the ratio scale. By definition, a ratio scale must have a well-defined zero point. For example, the dollar sales figure for Tide detergent in a particular store over one month period is a ratio scale. When comparing year to year sales, it is meaningful to state the current sales are a certain percentage higher than sales a year ago. In marketing research, consumers are sometimes asked to allocate 100 points across N brands in proportion to preference. This type of question, called a *constant sum scale*, can be considered a ratio scale of brand preference.

Although these distinctions may seem technical, they are important in determining which statistical analyses are appropriate for itemized rating scales. As noted earlier, itemized rating scales clearly have ordinal properties. However, it is common practice in marketing research to analyze the scales as if they have interval scale properties. In fact, itemized rating scales are also called *summated rating scales* because researchers will often compute an attitude score for each consumer by summing over a number of rating scales.

This ordinal versus interval measurement controversy has never been completely resolved. Early researchers in attitude measurement assumed that itemized rating scales provide interval level measurement, and felt justified in

analyzing the scales using statistical tools such as factor analysis. More recent developments in the psychometric literature, in particular *latent variable models* (Skrondal and Rabe–Hesketh, 2004), provide evidence that itemized rating scales are best regarded as ordinal scales, and should be analyzed with specialized procedures. We return to this controversy later in this article.

Itemized Rating Scales

The discussion so far has suggested that itemized rating scales have at least ordinal properties. Over the years, researchers have developed numerous scales having this property. Here, we briefly discuss three major types of itemized rating scales: *Likert*, *semantic differential*, and *Stapel*. These scales, shown in Table 2, are among the most commonly used scales in marketing research.

Likert scales. Likert scales ask the consumer to indicate his or her relative degree of agreement with respect to declarative statements. For example, we could write the statement "Wal–Mart has great service" and ask consumers to respond on a 5-point scale, where 1 means "Strongly Disagree" and 5 means "Strongly Agree." Likert scales are widely used because they can be adapted to studies of product positioning, consumer lifestyle, brand preference, customer satisfaction, and attribute importance. Likert scales (and, indeed, all the scales in Table 2) are *noncomparative* attitude measures: the consumer rates one entity at a time (*see* NONCOMPARATIVE SCALING TECHNIQUE). If the researcher wishes to compare attitudes across entities, the consumer must be asked to rate the various entities separately.

An interesting application of the Likert scale is *quadrant analysis*. In this analysis, consumers are presented with two sets of Likert scales, each scale with K possible responses. The first set asks the consumer the relative importance of various product attributes in selecting the product (1 = Unimportant to K = Important). The second set asks the consumer whether he or she believes that their usual brand has these same attributes (1 = Strongly Disagree to K = Strongly Agree). The responses on the belief scales are then plotted against the corresponding response on the importance scales. Because a large sample of consumers is surveyed, it is conventional to use the percentage of consumers using the top two rating categories (called *top–box responses*) as the plotting point.

An example of quadrant analysis for detergent brands is shown in Figure 1. Note that detergent attribute beliefs are positively associated with

Table 2 Itemized rating scales.

Type of Scale	*Example of Scale*				
	Wal–Mart provides great customer service.				
Likert	1	2	3	4	5
	Strongly disagree				Strongly agree
	Please rate Wal–Mart on the following scale:				
Semantic Differential	Great service	I------I------I------I------I------I			Poor service
	Please rate Wal–Mart on the following scale:				
			+3		
			+2		
			+1		
Stapel			Service		
			–1		
			–2		
			–3		

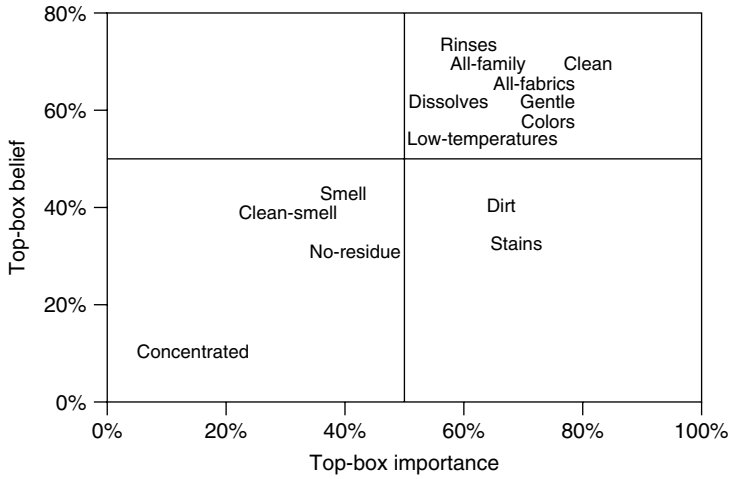


Figure 1 Quadrant analysis.

detergent attribute importance. This is owing to the fact that brand competition forces manufacturers to produce products with attributes desired by consumers. To create quadrants, each axis of the graph is split at some point – in this case, the 50% mark. The lower right quadrant (high importance, low belief) indicates attributes that manufactures should target for new-product development. In this case, consumers indicate that they value the ability of detergents to remove dirt and stains, but find their current brands lacking on these attributes.

Several practical considerations are associated with Likert scales. One of the first decisions is determining K, the number of scale points. Numerous psychometric studies suggest that more scale points are better in terms of information recovery, but only up to about seven points. Additional scale points do not provide substantially more information about the underlying attitude, and run the risk of confusing the consumer. A related question is whether the scale should have an odd or even number of points. An odd number allows for a middle (neutral) category that can be used if the consumer is ambivalent. In contrast, an even number of points force the consumer to choose one end of the scale or the other. A final question is whether or not the consumer should be allowed to opt out of responding to a Likert scale by selecting "Don't Know." Failure to supply a "Don't Know" option can lead to increased measurement error by uninformed consumers. However, allowing "Don't Know" creates missing data problems in the subsequent statistical analysis. Marketing research consultants tend to favor an odd number of response categories without an option for a "Don't Know" response. These choices, however, should be on the basis of the needs of the particular researcher.

Semantic differential and Stapel scales. The remaining two scales are variants of each other. The semantic differential scale is a bipolar scale for which the consumer rates an entity on some characteristic. For example, we could ask consumers to rate Wal–Mart along a 5-point scale, where 1 indicates "Great Service" and 5 indicates "Poor Service." The words on the endpoints are intended to convey the fact that the researcher is interested in the position of Wal-Mart on a continuum of service quality. In practice, problems arise if the consumer does not view the two endpoints as opposites.

As shown in Table 2, the Stapel scale is a semantic differential scale with an even number of response categories and the endpoints not labeled. The Stapel scale was developed, in part, to escape the problem of endpoint definition and to allow easier survey administration over the phone. Hawkins, Albaum, and Best (1974)

provide empirical evidence that the semantic differential scale and the Stapel scale can provide comparable information when the number of scale points in the two administrations is the same. The principal challenge of the Stapel scale is developing labels (such as "Service" in Table 2) that accurately convey the attribute that the researcher wishes to measure.

Scale Development

Often, the objective of the research study is to develop a measure of some sort of consumer characteristic (such as deal proneness) or consumer perception of some product characteristic (such as service quality of a fast food restaurant). To improve the accuracy of the attitude measure, researchers often ask the same question in a variety of different ways, and then average the answers to provide a final score. The process of selecting the items to be included in a composite scale is known as *scale development*.

Item generation. The first step is the generation of items to include in the survey. Here, the researcher must make a key decision. Is the objective of the research *person-centered* measurement or *object-centered* measurement? In the context of marketing research, *person-centered measurement* is defined as the development of an attitude measure for each consumer in the study. As such, the researcher will need to ask each consumer a number of related questions. In contrast, *object-centered measurement* is defined as the measurement of a characteristic of a product from a market-level perspective. For an object-oriented measurement application, it suffices to ask each consumer one question, and then to average the answers across all consumers in the study. To avoid confusion in the remainder of this article, we assume that the researcher is primarily interested in person-oriented measurement.

Item generation is a creative task that requires the researcher to understand the meaning of the attitude under study. Using a *deductive* (top-down) approach, the researcher would carefully define the construct under study and then attempt to create items that measure different aspects of the construct. Scales might be based upon previous research using the same construct. In contrast, using an *inductive* (bottom-up) approach, the researcher might discuss the topic with focus group consumers, and then use the focus group transcript as a basis for item generation.

Most scales are constructed using both deductive and inductive reasoning. Consider the deal proneness scale in Table 3, taken from a recent study by Wirtz and Chew (2002). The six items are all Likert items, having a 7-point scale from 1 (strongly disagree) to 7 (strongly agree). Because deal proneness is a relatively straightforward attitude, it is possible for a researcher to simply generate these types of items. However, talking to consumers in a focus group setting could provide additional insights that might suggest other items for inclusion. It should be understood that the six scales in Table 3 are just a subset of the original set of items generated in the research process. As we explain next, it is

Table 3 Deal proneness scale items.

1. Redeeming coupons and/or taking advantage of promotion deals make me feel good.
2. When I use coupons and/or take advantage of a promotional deal, I feel that I am getting a good deal.
3. I enjoy using coupons and/or taking advantage of promotional deals, regardless of the amount I can save by doing so.
4. I am more likely to buy brands or patronize service firms that have promotional deals.
5. Coupons and promotional deals have caused me to buy products and/or services I normally would not buy.
6. Beyond the money I save, redeeming coupons and taking advantage of promotional deals give me sense of joy.

All items are measured using 7-point Likert scales (1 = strongly disagree, 7 = strongly agree). Conceptual and methodological details concerning the development of this scale can be found in Wirtz and Chew (2002).

essential to generate more items than are really needed, in order to select only those items that best measure the attitude under study.

Scale construction. Scale construction is a process of evaluating the properties of an attitude scale using data collected from a representative sample of consumers. The psychometric theory of mental test scores (Lord and Novick, 1968) suggests two key considerations: *reliability* and *validity* (*see also* RESEARCH RELIABILITY AND VALIDITY). Simply put, a scale is reliable if the measurement error of the scale is small. A scale is valid, if it actually measures the underlying attitude that the researcher intends the scale to measure. It can be shown that high reliability is a necessary condition for high validity. However, it is entirely possible to have a highly reliable scale that is completely invalid.

The usual procedure is to generate a large number of items, construct a questionnaire based upon these items, and then survey a *representative sample* of consumers. The sample should be drawn from the same type of consumers who will be studied after the scale is finalized. For example, college students should not be used during scale construction if the scale is intended for use among middle-aged consumers. To avoid response biases, some items should be reverse coded, so that high agreement (on a Likert scale) or positive affect (on a semantic differential scale) is not always on the right-hand side of the scale (with an associated high scale number). Subject to limitations of respondent fatigue, the larger the list of potential scale items, the better.

Once the survey is administered, the researcher then must do a series of technical checks, both to weed out unacceptable items, and to ensure that the scale represents an unidimensional construct. It is here that the ordinal versus interval measurement controversy noted earlier becomes important. The usual practice is to simply assume interval level measurement of the individual scale items, and then to employ regression-based statistical tools to analyze the properties of the proposed attitude scale. Later in this article, we briefly discuss alternative approaches that rely only upon the ordinal properties of itemized rating scales.

At this point, the researcher is interested in two issues. The first is whether or not all items measure the same underlying attitude. This can be examined by estimating a correlation matrix for all items (using the survey responses) and then analyzing the correlation matrix using *exploratory factor analysis* (*see* EXPLORATORY FACTOR ANALYSIS). It is beyond the scope of this article to adequately explain the factor analytic model in any detail. Intuitively, factor analysis is a way of exploring the structure of a correlation matrix. The approach assumes that the estimated correlations arise from two sources: underlying attitudes that the consumer uses answer the scale items on the questionnaire, and random error because of the measurement process. The factor analysis model first adjusts the correlations for the reliability of individual scale items, and then develops an eigenvalue decomposition of the adjusted correlation matrix. If a one factor solution emerges, then the researcher can be confident that all items measure the same attitude. Usually, a one factor solution does not fit the correlation matrix. In this case, the researcher deletes unacceptable scale items until a one factor solution emerges. This is the basis of the general recommendation that the researcher generates a large set of items for consideration in scale construction.

The second issue is the issue of reliability. Formally, *reliability* is defined as the proportion of the observed variance in the scale rating that is owing to the true score (i.e., the underlying attitude the researcher is attempting to measure). Reliability can be measured directly by having the same group of consumers take the survey twice, and then correlate the rating selected for the same item across the two occasions. This procedure is both difficult to administer and assumes that consumers do not remember their prior ratings. Instead, the usual procedure is to examine the variances and intercorrelations among the items of the proposed attitude scale and then compute a measure of reliability known as *Cronbach's alpha* (Eagly and Chaiken, 1993).

Generally speaking, an alpha of 0.70 or higher is considered acceptable in scale construction. Adding more items to a scale increases reliability. However, psychometric theory indicates that attitude scales having 5 to 7 scale items are sufficiently accurate for most research purposes, as long as the Cronbach's alpha cutoff is attained. Wirtz and Chew (2002) report that the 6-item

Table 4 Hypothetical multitrait–multimethod matrix.

		Semantic Differential			*Stapel*		
		S	E	R	S	E	R
	S	(0.93)					
Semantic differential	E	0.47	(0.90)				
	R	0.21	0.35	(0.91)			
	S	[0.59]	0.23	0.21	(0.80)		
Stapel	E	0.20	[0.60]	0.13	0.15	(0.87)	
	R	0.22	0.24	[0.68]	0.33	0.10	(0.83)

scale shown in Table 3 has a Cronbach's alpha value of 0.73.) It is important to understand that a high Cronbach's alpha does not necessarily imply that the items form an unidimensional scale. Factor analysis is always necessary to verify scale dimensionality.

Scale evaluation. The final step in scale development is validity. As noted earlier, a valid scale measures the actual attitude that the scale was intended to measure. A variety of different ways of assessing validity exist. The two most important are *convergent validity* and *discriminant validity*. Convergent validity is the idea that two scales that purport to measure the same attitude ought to yield scores that are highly correlated. Discriminant validity is the opposite notion: scales which purport to measure different (unrelated) attitudes ought to yield scores that are not highly correlated. As is apparent from this description, validity requires the researcher to have a good theoretical understanding of the relationship among the different consumer attitudes being measured.

One of the most stringent tests of validity is the multitrait–multimethod (MTMM) analysis developed by Campbell and Fiske (1959). To understand the approach, consider the following hypothetical scenario. Suppose that a researcher develops two versions of a multiple-item scale to measure consumer perceptions of brand personality: the attribution of human personality traits to brand names. Three scales are created, for seriousness (S), excitement (E), and ruggedness (R). The researcher decides to develop one measure of brand personality using semantic differential scales, and another using Stapel scales. Two surveys are done, one month apart, during a period of time in which the advertising campaigns in the market undergo no changes in message. The same consumers are surveyed at both time points. The goal of the researcher is to validate the three personality scales.

A hypothetical MTMM matrix for this scenario is shown in Table 4. The matrix is symmetric with only the bottom half shown. Along the diagonal (in parentheses) is the Cronbach's alpha reliability measure for each type of scale. In this case, the semantic differential scale apparently has less measurement error than the Stapel scale. The correlations in brackets [x] are measures of convergent validity: the correlation between two scales attempting to measure the same aspect of brand personality. The remaining correlations reflect method variance: the observation that different attitude measurements constructed using the same type of itemized rating scales will be positively correlated. The overall pattern in the matrix suggests that semantic differential and Stapel scales for the same personality characteristic measure the same attitude (convergent validity) and that three brand personalities are distinct attitudes (discriminant validity).

Another type of validity analysis concerns marketing research conducted in various countries (such as the member countries of the European Union). Although itemized rating scales can always be used, survey items and respondent instructions must be translated into different languages. In this context, there are obvious questions about whether attitudes measured in one country (say, Spain) can be compared to supposedly identical scales (except for language) in another country (say, Italy). Steenkamp and Baumgartner (1998), using the statistical tools

of *confirmatory factor analysis* (Brown, 2006), show how to evaluate whether scales in different countries measure the same attitudes. Assuming that the scales pass this test, they further show how to adjust the scales so that meaningful cross-country comparisons can be made.

Latent Variable Models

As noted earlier, one of the running controversies in the attitude research community is whether or not itemized rating scales can be considered as interval level measurement. Advances in psychometrics in the last 40 years allow researchers to develop continuous measures of attitudes while assuming that rating scales only have ordinal properties.

The general approach is called *latent variable modeling* (Skrondal and Rabe-Hesketh, 2004). Simply put, a latent variable is a random variable whose outcomes cannot be directly observed. However, these outcomes impact observed variables in a known way. In psychometrics, latent variables are found in *latent trait theory* (LTT), a way of conceptualizing how people respond to rating scale items (Lord and Novick, 1968). In the context of itemized rating scales, LTT assumes an unidimensional attitude continuum in which consumers have an attitude score at a specific location, and the itemized rating scale imposes cut points along the same continuum. LTT assumes that the consumer attempts to report the interval along the continuum in which his or her attitude lies, but the reporting is subject to random errors with a known distribution.

Researchers working with marketing survey data have demonstrated that the LTT technology allows the simultaneous calibration of rating scale cut points and consumer attitudes (Balasubramanian and Kamakura, 1989). This work has interesting implications for optimal scale construction, including the idea that rating scales can be optimally tuned to each individual by selecting only those items that maximize the accuracy of the attitude measure generated by the scale. Such *tailored tests* are possible if respondents are seated in front of a computer terminal.

One of the most interesting recent developments in scale construction is the *partial least squares* (PLS) algorithm (Geladi and Kowalski, 1986) applied to latent variable systems. Traditionally, attitude scales are constructed without regard to their ability to predict other attitude scales. Indeed, the entire literature on scale validity (discussed earlier) is based upon the notion that scales should be constructed in a logically rigorous fashion and then validated by computing correlations with other independently generated scales. The PLS algorithm makes it possible to scale two sets of rating scales simultaneously in such a way that the correlations among the underlying latent attitude scales are maximized.

This approach is the methodological foundation of the American customer satisfaction index (ACSI), a satisfaction measurement system developed at the University of Michigan (Fornell *et al.*, 1996). Data collection for the ACSI consists of consumer surveys tapping aspects of customer satisfaction in a variety of goods and services, all done using 10-point Likert scales. The ACSI statistical analysis yields six continuous attitude measures (perceived quality, customer expectations, perceived value, customer complaints, customer loyalty, and overall customer satisfaction). Conceptually, the ACSI methodology creates attitude scales with known validity and reliability characteristics, while only assuming that the Likert scale inputs have ordinal measurement properties.

Summary

Itemized rating scales are ordinal scales used to measure consumer attitudes. Three types of scales are commonly used in marketing research: Likert, semantic differential, and Stapel. Likert scales ask consumers to agree or disagree to declarative statements. Semantic differential and Stapel scales ask consumers to rate entities along an ordered continuum. Attitude scales in marketing research are usually created as multiitem scales, which are analyzed using regression and factor analytic methodologies to examine reliability and validity. Newer analysis methods based upon latent variable models assume that itemized rating scales only have ordinal measurement properties, but allow the construction of continuous measures of consumer attitude.

Bibliography

Balasubramanian, S. and Kamakura, W.A. (1989) Measuring consumer attitudes towards the marketplace with tailored interviews. *Journal of Marketing Research*, **26**, 311–328.

Bishop, Y., Fienberg, S.E., and Holland, P. (1977) *Discrete Multivariate Analysis: Theory and Practice.*, MIT Press, Boston.

Brown, T.A. (2006) *Confirmatory Factor Analysis for Applied Research*, The Guilford Press, New York.

Campbell, D.T. and Fiske, D.W. (1959) Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, **56**, 81–105.

Eagly, A.H. and Chaiken, S. (1993) *The Psychology of Attitudes*, Harcourt Brace College Publishers, Fort Worth.

Fornell, C., Johnson, M.D., Anderson, E.W. *et al*. (1996) The american customer satisfaction index: nature, purpose and findings. *Journal of Marketing*, **60**, 7–18.

Geladi, P. and Kowalski, B.R. (1986) Partial least squares regression: a tutorial. *Analytica Chimica Acta*, **185**, 1–17.

Hawkins, D.I., Albaum, G., and Best, R. (1974) Stapel scale or semantic differential in marketing research? *Journal of Marketing Research*, **11**, 318–322.

Lord, F.M. and Novick, M.R. (1968) *Statistical Theories of Mental Test Scores.*, Addison-Wesley Publishing, Reading.

Skrondal, A. and Rabe-Hesketh, S. (2004) *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*, Chapman and Hall/CRC, Boca Raton.

Steenkamp, J.-B. and Baumgartner, H. (1998) Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, **25**, 78–90.

Stevens, S.S. (1946) On the theory of scales of measurement. *Science*, **103**, 677–680.

Wirtz, J. and Chew, P. (2002) The effects of incentives, deal proneness, satisfaction and tie strength on word-of-mouth behavior. *International Journal of Service Industry Management*, **13**, 141–162.

observation methods

Robert V. Kozinets

Development and Definition of Observation Methods in Marketing Research

For well over a half century, marketers have been interested in understanding the bases of brand loyalty. Marketing researchers have responded to this need by developing methods and techniques designed to illuminate the reasons that consumers show their loyalty to particular brands in marketplaces that are, for the most part, distinguished by their lack of differentiation. In the period from the 1930s to 1960s, motivation research was an extremely important set of techniques used by the marketers at major corporations to give them a competitive edge for differentiation (*see* MOTIVATION RESEARCH). These techniques were built around forming deep understandings of the way that consumers feel about and interact with products and services. Observation methods grew out of the recognition that gaining an understanding about the reasons for consumers' behaviors was critically important to competitive and proactive marketing management. To gain this realistic understanding, marketing research required an orientation that considered product and brand usage in a natural, holistic, and cultural context (Sunderland and Denny, 2007). This meant observing actual product purchase and consumption where it occurred (*see* ETHNOGRAPIC RESEARCH). Observation methods were specially adapted to this reality (Angrosino, 2007).

Observation methods are defined as qualitative marketing research methods in which researchers view, record, and then analyze the manifest actions of consumers as they engage in some market-related activity. The range of observations and data is considerable. Observation methods would cover the mass observations and measurements of large groups of anonymous people (*see* PERSONAL OBSERVATION). They would also encompass the placing of cameras in public places or in private areas. Analyses of the data from observation methods would also possess a considerable variety of forms, ranging from automatic counters calculating basic behaviors to extremely intimate investigations where researchers intrusively live with consumers in order to carefully observe and record their behaviors.

Origins of Observation Methods

Observational research and its methods originated in the anthropological method of ethnography (Angrosino, 2007). *Ethnography* is an anthropological method that studies members of a culture in depth through the techniques of participant observation (*see* ETHNOGRAPIC RESEARCH). Ethnographers conduct naturalistic research – research that is done in a naturally occurring context. Laboratory experiments and pen-and-pencil surveys and tests would not be considered naturalistic (*see* VALIDITY IN EXPERIMENTATION). An ethnographer's insights are gathered from "in-the-field" observations of people as they go about their daily lives (Sunderland and Denny, 2007). Although ethnography was developed originally for studies of distant, "primitive" tribes, it was applied to urban cultures in the early decades of the twentieth century. By the second half of the twentieth century, the advantages of ethnography had been recognized by the marketing and marketing research communities, and dozens of specialized firms and practitioners turned their unique observational skills toward observation of consumers in their native habitat (Angrosino, 2007). Observation methods in marketing research thus specialize, adapt, and then apply the observational techniques developed by anthropologists.

As observation-method practitioner Bill Abrams (2000, p. xxi) puts it, "there is no research tool like observational research for providing personal, up-close insight. No other research methodology allows you to go through your consumer's pantry, stand there while he/she flosses his teeth, or watch while he/she suffers and gets relief from an upset stomach or be there when a new computer owner gives up trying to get it to work with his/her printer at one in the morning."

This level of closeness to the consumer offers invaluable intelligence for marketers. It is difficult, but extremely valuable to understand how

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

consumers actually behave in certain situations, rather than simply ask them about how they behave or intend to behave (*see* PERSONAL OBSERVATION). As we know from many statistical studies, consumers' memories and recollections about their behavior are unreliable, as are measurements of their intentions (*see* RESEARCH RELIABILITY AND VALIDITY). If, for example, we were to ask the average large American family how many breakfast cereals are in their kitchen cupboard, or how many different bottles and brands of shampoo are in their bathrooms, we would be extremely unlikely to get an accurate answer. Yet, marketing researchers commonly ask such questions in their surveys and in focus groups. Observation methods offer an alternative to these after-the-fact methods of data collection. Instead of asking these questions, the researcher gains access to the consumers' homes and actually burrows through their kitchen cabinets, bathroom counters, bedroom closets, garage, and even garbage cans looking for the details and the clues about what products actually have been purchased and are in use, and what behaviors actually occur (Sunderland and Denny, 2007). With important corporate decisions on the line, observation methods can be an extremely powerful tool for revealing actual behavior.

Procedures and Examples of Observation Methods

Consider the needs of a hypothetical fast food company seeking marketing research about traffic patterns. Let us say this company, Joe's Burgers, is trying to make a decision about which location to lease to open a new restaurant. Observation methods could be brought to help make this decision. An example would be to hire someone who simply stood with a counter and stopwatch and manually counted the number of cars that passed by particular intersections of the city. Moving from one intersection to another and sampling at several different times of the day would yield comparisons that could help the fast food company choose the right location for its business. Another simple example would be to conduct market research at a cross-border shopping mall to understand how many consumers drive across the border to shop there. A survey researcher might consider asking a sample of 400 or more randomly chosen shoppers about their origins. A researcher using observation methods, however, would simply head to the parking lot. There, the researcher would note the license plates, using the actual cars in the lot to determine in which country the shoppers reside.

Observation methods can also involve the installation of cameras, or the use of existing cameras – such as the many closed-circuit television (CCTV) cameras in metropolitan areas. In the case of installed cameras, consumer homes that are of interest to the marketing researchers are chosen. The consumers are compensated for installing a camera in a particular room (let us say it is the kitchen). The camera stays on and records the consumers' behavior 24 hours a day. The resulting videotapes are analyzed carefully, sometimes even frame-by-frame for the intelligence and understanding they can provide of consumers' habits (Angrosino, 2007).

At their most intimate and most revealing, observation methods can involve an in-home interview and observational session or sessions. Observational interviews (also called *in-home interviews* or *ethnographic interviews*) will take place in a naturalistic setting, namely, in any setting that is familiar to the consumer (*see* FOCUS GROUPS AND DEPTH INTERVIEWS). They do not usually take place in the presence of other consumers or behind one-way glass, as with focus groups. Observations of product use and the consumer interview take place in the location where the product is usually bought and used. In those natural settings, such as consumers' homes, favorite stores of the consumer, or consumers' workplaces, the consumer tends to relax and generally finds it much easier to do and say what they usually do and say (Sunderland and Denny, 2007).

So, for example, in a study of toothpaste, the observation methods interview would take place in the consumer's own bathroom. It would also take place at the usual time when the consumer uses the product – in this case, first thing in the morning or the last thing before bed at night. After gaining permission from the consumer, and ensuring consent, the interviewer would "tag along" while the consumer gets ready for bed, or wakes up (usually this requires financial compensation). During the interview, the researcher might investigate the area and related

areas in the home. For a toothpaste inquiry, the researcher might look around the bathroom and examine the toothbrushes, mouthwash, and other oral hygiene products that might be on the counter. The researcher might also investigate the cupboards where extra tubes of toothpaste and toothbrushes are stored.

The interview is recorded, oftentimes with an audio recorder. The researcher usually keeps a notepad nearby and jots down notes for future reference as well as to immediately follow up on interesting phrases or topics raised by participants during the interview. Still pictures are often taken and, increasingly, video cameras are used to capture the details as well as simply to record the interaction. Procedurally, Abrams (2000) says that it is very important for the interviewer to maintain a relaxed, laid-back, and easygoing attitude during the interview. At the same time, the interviewer also uses a thoroughly designed and precise interview guide to ensure that all questions are being answered. As in depth and ethnographic interviews, the researcher wants to gain the trust of the consumer and engage in genuine rapport.

Observation research can be conducted as a team. Usually, a small team of marketing and research practitioners enter the homes of a small number of research participants. Usually, there are between 10 and 20 participants scheduled for such a study. The team carefully observes, films, records, and take notes of the individuals and/or the family going about their consumption or shopping behaviors, while trying to stay as unobtrusive as possible. The team then regroups and debrief together in an office space, a restaurant, or, perhaps, a hotel room. They go over their notes and make plans for the next session. At the end of this "blitzkrieg" of "home invasion" research, which usually lasts three to five days, the team will write their findings and present recommendations for action to the management (Abrams, 2000).

With a skilled researcher, consumers will start to relax, act naturally, and disclose their opinions and preferences within just a few minutes. Oftentimes, family pressures and politics are discussed as motivating factors behind many decisions. The reseacher rapidly becomes privy to confidential matters, such as husbands' deep-seated need for their wives' approval, wives' deep-seated need for their husbands' approval, parents' deep-seated need for their children's approval, and even some consumers' "dirty little secrets" that they do not floss/brush/comb/clean nearly as often as people think they do!

Another common observation method used to audit and research retailers is the use of mystery shopping or the "mystery consumer." In the mystery shopping methodology, researchers enter a retail environment posing as ordinary consumers. They gather specific – often highly specific – information about the quality of the retail service provided, as well as particulars of product choice, display, lighting, and other details. After leaving the retail environment, these hired professional observers report their findings either to other researchers who aggregate the data or to the final client.

Videography

The "videography" technique is gaining acceptance and popularity (Belk and Kozinets, 2005). Videography allows marketers to access, analyze, and research the actual actions of consumers as they go about their consumption in any naturalistic context (*see* PERSONAL OBSERVATION). The expenses associated with producing and distributing broadcast-quality video have drastically fallen as digital technology has advanced over the past decade. These developments have led marketing researchers to enter the domain of filmmaking once reserved for documentary journalists.

The most common use of videography is the videotaped individual or group interview, conducted either in a research facility or in a field setting (Belk and Kozinets, 2005). Videotaped interviews offer the powerful advantage of capturing body language, distance, motion, and other temporal and spatial dimensions of human behavior and social meaning. The second most common use of video is as a form of observation method: using the video camera to record naturalistic observations. In recording naturalistic observations, the videographer is more interested in capturing what people actually do rather than what they say about what they do (Kozinets and Belk, 2006).

Sometimes, a camera person accompanies the researcher, usually with a small camera and a lavalier or a wireless microphone. Increasingly common is the videographic technique where the interviewer is also the cameraperson. The filming techniques that are usually used in observational research originate from a school of film known as *cinéma vérité*, which was developed by documentary filmmakers in the United States and France in the late 1950s. The theory behind *cinéma vérité* is for the filmmakers to become as unobtrusive as possible. In its application as a form of consumer research, the idea is for consumers to get used to the presence of the videographer. As the videographer starts to naturally blend in to the context of the consumers, he or she can start to capture surprising revelations as well as genuinely mundane aspects of daily life.

Variants of videography based on naturalistic observation are increasingly common. Sometimes, videographers provide these research participants with a camera and ask them to directly capture and videotape their usage of a product or a service. This autobiographical technique has distinct advantages over researcher-conducted video observation. It is considerably less of an intrusion into the participant's life and is also less directed by researcher motives and needs, perhaps allowing new insights to come to the fore (Kozinets and Belk, 2006).

Unobtrusive observations caught on video, such as store camera data, are also commonly used as a form of videographic observation method. Sometimes, these observations can use concealed video equipment, a practice that must be used cautiously as it can be unethical. Like all observation and qualitative research, videographic data analysis follows the basic principles of interpretive analysis, from coding, to grounded theory construction, comparative case analysis, and hermeneutic circling. Video insights can be generated in a holistic fashion, or can occur through a bottom-up process that codes and classifies video capture on a frame-by-frame, event-by-event, or moment-by-moment basis.

Videography is a compelling observation method with many advantages, which is why it has become an increasingly frequent accompaniment to traditional marketing research interviews and ethnography. Videography captures the moment and allows it to be analyzed by researchers. Not only this, but it also can be made portable, shared with other researchers, and incorporated into training materials and research presentations. With audiovisual research, the marketing research is able to capture more of the subtle temporal, social, and spatial dimensions of human behavior, including the intricacies of human body language. The resulting research presentation to management is often emotional, vibrant, and humanizing in a way that slide presentations and written reports simply cannot match (Belk and Kozinets, 2005; Kozinets and Belk, 2006; Sunderland and Denny, 2007). Marketers watching a well-done videography often gain a sense that they have actually spent time with their consumers, experiencing their behaviors alongside them.

Marketing and Retail Research Uses of Observation Methods

Observation methods have been used to develop insights for strengthening the brand differentiation of a variety of products and services. In a competitive marketplace, even small differences in perception or/and function can be critical to gaining market share. It was observational research conducted by a cereal manufacturer that led to a repositioning of a common O-shaped cereal as a snack food for infants. In observations of breakfast cereal use, observational researchers discovered that mothers would often pack a plastic bag of these oval-shaped oat cereals to take it "on-the-go" as finger food to feed their young children out of the home. The finding was new to the cereal manufacturer. Focus groups, surveys, and personal interviews had not revealed it. In response, the manufacturer began marketing the cereal as a finger food for young children, promoting it in giveaways and children's books, and offering convenient plastic containers with prominent branding on them to replace mothers' inconvenient and flimsy plastic bags. The result was incremental sales for the manufacturer in an extremely competitive marketplace, better targeting at a new segment, and stronger branding for the cereal that has continued to the present day.

Observational research has also been found to be extremely helpful in generating insights and identifying areas of untapped opportunity for new-product development. By watching how consumers use – and misuse – existing products, and by speaking to them during their product use about their frustrations, unfulfilled needs, and dream or fantasy products, valuable raw ideational material is generated (*see* EXPLORATORY RESEARCH). These consumer insights can then be translated into working prototypes, tested on representative samples, subject to cost–benefit analyses, and eventually turned into successful new products.

Consumers in observational studies often generate valuable ideas spontaneously. Observational studies of how consumers "fix" the problems that they find in products often lead to new ideas for innovations (Sunderland and Denny, 2007). For example, observational research on how people used condiments at home revealed a common problem. The problem revealed itself once consumers got close to the bottom of a squeeze bottle of ketchup or mustard. At that point, the last 20% or so of the bottle was difficult to squeeze out. What do consumers do? They turned the bottles upside down, of course, developing various contraptions to balance them in their rotated positions. When engineers at the food company studied pictures of these almost-toppling bottles, they came up with a packaging solution to the problem. The solution was to design the bottles with a flat top so that, once enough of the condiment had been squeezed out of the bottle, it could be easily turned over and stored. The result was a new-product packaging innovation, competitive differentiation in a highly competitive category where steady gain was being made by private label brands, stronger consumer packaged goods branding, and a consumer problem solved, all courtesy of observation methods.

Observation methods also help retailers to understand the pain points and difficulties faced by shoppers as well as the many opportunities for improving consumers' shopping experiences. In a hidden camera study of a hardware store, the confusion on consumer faces and in their body language is almost comical. The consumers are faced with a relatively simple task such as matching the desired screw with its counterpart from the store. Yet with dozens of choices, and no organization or system for the shelving, the task was nearly impossible for most people. The observation methods study resulted in the hardware retailer designing a shelving system with clearly marked differences, enabling the average consumer to quickly and efficiently find the screw that they were looking for.

Consultant Paco Underhill (1999), the founder of retail marketing research agency Envirosell, relies almost exclusively on observation methods to generate insights for clients such as Wal-Mart, Saks Fifth Avenue, the Gap, Hallmark, and Starbucks. Underhill's research consists of careful tracking of what he calls "the science of shopping." "Essentially, trackers stealthily make their way through stores following shoppers and noting everything they do. Usually, a tracker begins by loitering inconspicuously near a store's entrance, waiting for a shopper to enter, at which point the 'track' starts. The tracker will stick with the unsuspecting individual (or individuals) as long as the shopper is in the store (excluding trips to the dressing room or the restroom) and will records on the track sheet virtually everything he or she does. Sometimes, when the store is large, trackers work in teams in order to be less intrusive" (Underhill, 1999, p. 13). In addition to counting, recording, and measuring the motions and behaviors of shoppers during their shopping trips, the trackers also record their own observations and insights in a set of field notes. These field notes are intended to capture the details of consumers' shopping behavior, making informed inferences based on the recorded observations (Underhill, 1999).

One of the principles discovered by Underhill and associates and recorded in his book about "the science of shopping" is something that he terms *the butt-brush effect* (Underhill, 1999, pp. 17–18). In a study that they conducted for Bloomingdale's in New York City, the observation researchers had a camera recording one of the main entrances. The camera also caught the next sales rack that was positioned on a main aisle near the entrance. As they reviewed the recordings, they noticed that shoppers approached the tie rack and stopped to look at the ties. However, because the aisle was near the entrance, people

who were heading out or in the store would bump the tie shoppers. After they were jostled once or twice in this manner, most of the shoppers would abandon their search for neckwear. Underhill's group concluded that shoppers did not like being touched, brushed, or bumped from behind and would cease their shopping in order to avoid it. This was true especially of women, but also of men. Upon further investigation with Bloomingdale's, it turned out that sales from the tie rack were lower than they had expected. Underhill's group hypothesized that "the butt-brush factor" was to blame. Bloomingdale's management moved the tie rack shortly after the presentation of results and a few weeks later sales from that rack apparently rose quickly and substantially. The results of this observation-method-based finding underscored the importance of positioning sales displays so as to allow freer movement and less unintentional bodily contact between people.

Another insight into consumer motivations and psychology was gained from retail use of observation methods. Underhill's group was conducting a supermarket study for a manufacturer of dog food. Observing the pet food aisle, the group noticed that dog treats tended to often be picked out by senior citizens and children. The group asked itself why this was happening. Their answer was that elderly people considered the dogs to be akin to their grandchildren, and wanted to spoil them. Similarly, children had fun feeding treats to dogs; it was almost like playing with a toy. Parents would buy dog treats for their playful children, giving in to their whims the same way they would indulge the children themselves with treats. The retailing implication came in regarding the placement of the dog treats. The pet treats were typically located near the very top of the supermarket shelves. In that location, children and elderly people often found it difficult to get at them. The advice coming out of this observational research was to move the treats where they could be more easily reached by older people as well as children. According to Underhill's (1999, p. 19) book, "the client did so, and sales went up overnight."

Technological Observation Methods and Trend Spotting

Internet and information communications researchers are among the pioneers of a new generation of observational research (Angrosino, 2007). Information monitors such as the Nielsen Ratings and TiVo have been working on observational technologies to unobtrusively measure people's media viewing habits. Tracking and following click stream data – which often links to individual IPs or Internet protocol addresses – is one sophisticated way to understand what people actually do online. A large number of social media companies, such as Google, Flickr, Amazon.com, and FaceBook have benefited from the ability of the Internet to track what people click on or actually "do." Technological tools applied to the online environment allow us to closely monitor and measure what people actually buy online, finding the patterns of online purchase or e-commerce. Trend-watching and trend-spotting companies are using new technologies to get more rapid feedback on the latest changing trends. Furthermore, the millions of conversations on the Internet have become an important source of data for marketing researchers using technologically enabled observation methods like data mining, thematic analysis, and netnography.

The Neilson Online measurement systems now reveal consumers' online viewing habits, for example, revealing the most popular streamed television shows and web sites. TiVo, the company that sells the digital video recording device and its related online service, is also pioneering methods of collecting data about television viewers from their product. In 2004, TiVo was the first in the primary source of information about how many people actually saved or "TiVoed" the image of Janet Jackson's exposed breast during SuperBowl XXXVIII. Although TiVo does not store any individuals' viewing records, and viewers can opt out of having their own personal data collected if they wish, TiVo receives a steady stream of information or data from its viewers, and it has begun selling the data to the television industry through various commercial research partnerships.

In another example of a technologically mediated use of observation methods, Look-look, a trend-forecasting firm has recruited thousands of young people between the ages of 14 and 30 to act as amateur trend spotters and observers. These consumers send in a constant stream of market-based information by instant messenger and pagers. The research firm conducts on-the-spot observational research designed to try to understand the trends and motivations of younger consumers. In this research, these hand-picked field correspondents also upload images from digital cameras that can record events such as parties, concerts, and sporting events. Information flow also goes both ways. Similar trend-spotting firms are constantly on the lookout, such as Boston-based BzzAgent, a word-of-mouth marketing firm that not only distributes products to its group of "influentials," but also monitors their opinions to stay on top of the latest trends.

With well over 100 million consumers as members of online communities globally, massive amounts of public interactive data exist on the Internet. For over a decade, marketing researchers have been devising methods to try to analyze and understand this information. When we consider the fact that consumer-to-consumer conversations are naturally occurring on the Internet, these computer-aided observational techniques offer a powerful combination that is both naturalistic and unobtrusive. Companies like Nielsen Buzzmetrics and MotiveQuest have pioneered sophisticated tools to analyze these online conversations. Their programs and analysts mine the Internet for the data, then code and analyze the data to help marketing managers understand the general complexion of Internet "chatter."

Another observation method that has developed as a direct outbranching of ethnography with the power of the Internet is called *netnography* (Kozinets, 2002). Netnography is ethnography adapted to the unique contingencies of the online environment: technological mediation, anonymity, textuality, automatic archiving of messages, and invisible public audiences. Netnography can potentially be conducted in a manner that is entirely unobtrusive. This sets netnography apart from other observational techniques such as in-home observation and ethnographies, as well as from more popular methods like focus groups, depth interviews, and surveys.

In terms of time expenditures, expending on recruitment, personal meetings, traveling, transcribing observations, conducting personal interviews, and writing handwritten field-note data, netnography is far less time consuming and elaborate than traditional, in-person observation methods such as in-home observational research interviews and ethnography. Netnography provides marketing researchers with a window into naturally occurring behaviors, such as community discussions about different products, services, lifestyles, or experiences. Because it is conducted using observations and interactional data collected from people in a context that is not designed or artificially fabricated by the researcher, netnography is less intrusive than focus groups and personal interviews. The type of information provided in a netnography and the ease of its collection means that netnographic research can be more timely than traditional, in-person surveys, interviews, and focus groups (Kozinets, 2002). As an observational technique of the naturalistic communities emerging and increasing on the Internet, netnography's potential is being recognized by global businesses and its popularity is growing. However, because of its cultural approach, netnographic data can be as difficult to interpret as ethnographic data. In many cases, netnography does not necessarily allow us access to what consumers actually do with particular brands, products, and services. It only allows us to "overhear" the naturalistic conversations that they have pertaining to them.

Limitations of Observation Methods

The limitations of observation methods tend to arise from their small sample sizes (*see* STATISTICAL APPROACHES TO DETERMINING SAMPLE SIZES). Because this is inductive research based upon the detailed analysis of very small groups of people, it is important for marketers to test these insights further to find out if they are idiosyncratic and unique, or broad and generalizable (*see* HYPOTHESIS TESTING RELATED TO DIFFERENCES – PARAMETRIC TESTS). If observation methods are appropriately analyzed and interpreted,

their conclusions will be carefully verified on broader samples using causal methods. These investments in verification should be made before other costly investments in production and marketing are made.

Because it tends to be time intensive, and to involve significant researcher skill, some observation research can be quite expensive to conduct. However, new technological forms of observation can be quite inexpensive to perform. Because observation methods now include such a vast diversity of forms, observation research no longer means sending sole researchers out into the field of consumers' homes to conduct weeks of grueling in-home interviews (*see* EXPLORATORY RESEARCH). Many companies are now finding that they can tailor an observational research plan to their specific needs.

In addition, observation methods provide personal, meaning-laden, symbolic, cultural-level data rather than statistical forms of data. Human actions and words require careful and skilled interpretation. Observation methods, like ethnography and many other methods (even quantitative ones) depend upon the skill of the researcher utilizing them (*see* ETHNOGRAPIC RESEARCH; EXPLORATORY RESEARCH; FOCUS GROUPS AND DEPTH INTERVIEWS; PERSONAL OBSERVATION). Some researchers might walk into a frat house filled with spray-painted beer bottles and see nothing but a mess. Others might see an opportunity to develop a new brand of beer that caters to a new, wild and free, do-it-yourself target segment, and their creative ethos. The skill of the observational researcher is critical to the interpretation, and can be the single, most important factor determining quality outcomes from the research. Nonetheless, in many cases, the raw data that is generated from observation methods offers something uniquely valuable: authentic observations of real people as they actually use products and services. Without the genuine data of real peoples' consumption generated by observation methods, true marketing insights and breakthroughs may be much harder to find.

Bibliography

Abrams, B. (2000) *The Observational Research Handbook: Understanding How Consumers Live With Your Product*, NTC, Lincolnwood.

Angrosino, M. (2007) *Doing Ethnographic and Observational Research*, Sage Publications, London.

Belk, R.W. and Kozinets, R.V. (2005) Videography in marketing and consumer research. *Qualitative Marketing Research*, **8** (2), 128–141.

Grossman, L. (2003) The Quest for Cool. Time Magazine (Aug 31) 2003.

Kozinets, R.V. (2002) The field behind the screen: using netnography for marketing research in online communities. *Journal of Marketing Research*, **39**, 61–72.

Kozinets, R.V. and Belk, R.W. (2006) Camcorder society: quality videography in consumer research, in *Handbook of Qualitative Research Methods in Marketing* (R.W. Belk), Edward Elgar Publishing, Northampton, pp. 335–344.

Sunderland, P.L. and Denny, R.M. (2007) *Doing Anthropology in Consumer Research*, Left Coast Books, Walnut Creek.

Underhill, P. (1999) *Why We Buy: The Science of Shopping*, Simon & Schuster, New York.

questionnaire design

Jagdish Agrawal

Introduction

Every day, managers are confronted with uncertainty regarding the future course of events and the possible outcomes of a decision. For example, a manager may need to react to new information provided by one of his salespeople that a major competitor has been offering deep discounts to its customers. Or, she/he may be confronted with the option to choose one of three different countries for a major investment to build a new plant in order to cut manufacturing costs. These uncertainties, also known as *problems*, require decisions. Some problems require an immediate decision based on whatever information is available at that time; others may require collecting new data from a variety of sources.

Conducting surveys with the help of a questionnaire is perhaps the most dominant method of collecting data for academic research, public institutions, and business organizations. A questionnaire affects both the quality and quantity of data. A poorly designed questionnaire will cause errors or bias in research findings rendering the results of the survey worthless for decision making. In addition, if the survey is not properly designed, it may not motivate respondents or it may discourage them from completing the survey resulting in the lack of adequate data to make valid inferences regarding the topics of research. Although it is perhaps impossible to design a perfect survey, we can avoid mistakes commonly made when designing a survey and improve both the quality and the quantity of data collected from the survey. The steps required to help achieve this objective are discussed in the following.

Identify the information needed. Before conducting any kind of survey, it is very important either to list all the questions that need to be answered through the survey or to develop hypotheses regarding possible causes or solutions to the problem(s). For example, let us assume that a manager of a supermarket is worried about dwindling store sales. On the basis of personal observations or customer complaints, she/he may, for example, hypothesize that declining sales are due to competing stores providing better services. This hypothesis may or may not be true, and the real reasons may lie elsewhere. We need to collect and analyze the relevant data to see whether this educated guess or hypothesis is true. This is very similar to a crime investigator developing hunches regarding the person(s) who may have committed the crime. In the absence of any hunches regarding the underlying reasons for the problem, the alternative is to develop a series of questions and seek answers to those questions by collecting data. Examples of such questions, called *research questions*, are "What do supermarket shoppers care about when shopping?" "How is my store different from the competing store with regard to the things shoppers care about?" It is possible to have both hypotheses and research questions at the same time. These hypotheses and research questions essentially decide the nature and scope of a research study. Assuming that the nature of the study requires survey research, specifying the research hypotheses and/or research questions helps to identify what information is needed from the survey, or any other data collection methods, as indicated in Table 1.

For practical research that is designed to solve the problem(s) or uncertainty faced by a manager, perhaps research questions may be more useful than developing hypotheses. A research question does not require any assumption regarding the source of the problem or possible solutions to that problem. However, research questions help the researcher to be focused on answering certain questions in a given study. For this reason, it is not advisable to make the scope of any one research study too broad – do not include too many research questions in one research study.

Select the appropriate survey method. Once we know what data are actually needed to reduce the uncertainty or "solve the problem," and assuming that we have decided to conduct a survey to collect the desired data, the next step in the research process is to decide the survey method that is best suited to collecting the desired data. There are two major methods of surveying:

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Hypotheses, research questions and data needed.

Hypotheses or Research Questions	*Hypotheses and/or Research Questions*	*Data Required to Test Hypotheses/Answer Research Questions*
Hypothesis	Competing store provides better customer services than the services we provide	Shoppers' perception of service quality at both stores
Research question 1	What do supermarket shoppers care about?	Relative importance of goods and services
Research question 2	How is my store different from the competing store regarding things shoppers care about?	Shoppers' perception of goods and services at both stores

1. Survey administered by an interviewer. Such a survey could be (i) in person or what is called a *face-to-face interview*, or (ii) over the phone with or without computer assistance.
2. Self-administered survey. This includes (i) mail surveys and (ii) online or Internet surveys.

A comparison of these four methods is presented in SURVEY RESEARCH. Basically, we need to consider a number of factors and choose the survey method that seems to be most suitable given those factors. These are listed below:

1. The characteristics of the target population help decide which method may be most suitable. For example, the education level may determine whether a face-to-face interview is needed to explain the questions and record the responses. Similarly, if we are conducting a survey of middle-level managers, perhaps an online survey will be the most appropriate way to reach them.
2. Cost of conducting surveys varies across survey methods. The choice of method may depend upon the budget available to conduct the survey. An online survey is likely to be the least expensive whereas the face-to-face interview is the most expensive.
3. The response time of a web-based survey is faster than that of the mail survey and the face-to-face interview in that order. If there is an urgent need for the data, either the online or the telephone survey would be more desirable than the other methods.
4. If there is a need for including visuals (graphics)/audio in the survey, the web-based survey and the face-to-face interview are better suited to achieving this objective than the other methods.

Select the type of questions. There are a variety of forms of questions from which to choose when designing a questionnaire. Every question in a survey is designed to measure what is called a *construct*. A construct is what we are trying to measure, and it could be anything – degree of satisfaction with the brand, the income of respondents, consumers' preference for certain features of a brand, the price consumers are willing to pay for a new brand, brand loyalty, consumers' reaction to the quality of service provided by a dealer, and so on. It is very important to have a clear understanding of what we are actually trying to measure. For example, let us suppose that to develop a demographic profile of consumers, we need to collect data on each consumer's income along with age, occupation, and so on. We need to define whether we need to have information regarding a respondent's income or household income? Income before tax or income after tax? Similarly, if we are trying to measure brand loyalty, we need to decide whether we are going to measure brand loyalty based on a consumer's attitude toward the brand or based on purchase behavior or both. Once we have a clear definition of what we are trying to measure, we could

choose to measure that construct or the variable by using different types of questions.

Open-ended versus closed-ended question. A closed-ended question is one for which there are a fixed number of responses. The respondent may be asked to choose only one, choose two, or choose all of the options available for the questions if applicable. For example, the number of options to choose from in the following question is fixed.

To which of the following magazines do you currently subscribe?

( ) Time
( ) Newsweek
( ) Business Week
( ) Fortune.

On the other hand, an open-ended question is one that does not provide a fixed number of options to choose from. In an open-ended question, the respondent is allowed to choose the answer in his own way rather than choosing an answer from the list provided by the researcher. Sometimes, it becomes necessary to include open-ended questions because there may be a situation where it may be difficult to know a priori the possible answers to a question. For example, if one wants to know what things the respondent likes about a city – let us say San Francisco – it is very hard to provide an exhaustive list of options. One can always restrict the answers to a few sentences or a few categories – "(Please give no more than three reasons for . . .).''

Although open-ended questions may be necessary and the detailed answers we get to such questions may be more valid and insightful than answers to a closed-ended question, one of the biggest challenges associated with such questions is the coding of answers. Open-ended questions have a tendency to invite lengthy and sometimes irrelevant responses from the respondent. Since each respondent is answering the question in his/her own words, It will require at least two people to code the responses to ensure an unbiased coding of responses for analysis and will likely take countless hours to complete the coding. Thus it is both labor intensive and costly to code such responses. One way of avoiding open-ended questions is to conduct a pilot test of such an open-ended question. Generate the frequencies of answers obtained through the pilot test and use the most frequently mentioned categories as options in a closed-ended question in the final questionnaire. An "other" category can be added to the options to capture the rest of the infrequently mentioned answers.

Yes/No question versus multiple-point scale. Although simple to ask and seemingly easy to answer, such dichotomous options are appropriate for situations where answers are quite definite. For example, a question such as "Do you smoke cigarettes?" or "Do you have access to the Internet at home?" can be answered with a "Yes" or "No" since they are asking for the truth rather than an opinion. However, if we are interested in human psychological constructs such as attitude, perception, beliefs, and intention to buy a product, providing a simple "Yes" or "No" option is not likely to capture a range of answers. For example, for a question such as "Do you like shopping?" or "Would you be interested in buying this product if it is available at $50?", a "Yes" or "No" option simply forces the respondent to choose from options that may not reflect a true intention. In this type of situation, providing a range of options (a five-point scale) such as those presented below is more likely to capture the possible range of likely responses than would a simple Yes/No dichotomous response as in Annexure I.

Thus, while a Yes/No response category may be appropriate for questions related to

Annexure I

Definitely would buy (1)	Probably would buy (2)	Might/might not buy (3)	Probably would not buy (4)	Definitely would not buy (5)

Annexure II

	Highly Disagree				*Highly Agree*
What I save by using coupons is not worth my time	1	2	3	4	5
Using coupons saves me a lot of money	1	2	3	4	5
An item on sale is likely to be of poor quality	1	2	3	4	5
It does not make sense to buy an item at its regular price	1	2	3	4	5
People who use coupons are smart shoppers	1	2	3	4	5
Every penny I save by using coupons helps my family	1	2	3	4	5

actual behavior or questions asking for facts, providing multiple options is necessary to measure constructs associated with people's beliefs, perception, attitude, intention, and so on. Researchers have suggested different types of scales to measure such psychological constructs. They are described in detail in ITEMIZED RATING SCALES (LIKERT, SEMANTIC DIFFERENTIAL, AND STAPEL) of this encyclopedia.

Number of points on a scale. This also brings us to the issue of how many points should be used in a scale. This is an issue researchers have studied and debated. Researchers have shown that five to seven points should be enough to capture consumers' responses to the psychological and social constructs that we tend to measure very often in business surveys using rating scales. A scale with only three points may be too few, and similarly a scale with nine points may be too many.

Availability of a neutral option. Another related issue is the question of whether to provide the midpoint (odd or even number of points on a scale). Although providing the midpoint (might/might not buy option in the previous example) allows the respondent to choose "neutral" if she/he is unsure about the issue, this also allows for an easy answer to a question if the respondent does not want to spend time thinking about how to respond. One way of deciding whether to include the neutral option is to discuss this issue with the respondents of the pilot survey during the debriefing session.

Using multiple questions/scales to measure a construct. Sometimes, it may be necessary or advantageous to use multiple statements to measure a construct rather than asking only one question. For example, if we are interested in measuring consumers' price sensitivity, we may ask multiple questions in different ways. This allows us to check the consistency of the answers as well as capture different aspects of *deal proneness*. Asking multiple questions to measure a construct also helps to test the reliability of those items in measuring a construct. An example of using multiple items to measure price sensitivity or deal-proneness is provided in Annexure II. Another example is provided in Annexure III.

Other essays in this book describe a number of comparative (*see* COMPARATIVE SCALING TECHNIQUE) and noncomparative scales (*see* NONCOMPARATIVE SCALING TECHNIQUE) available to measure psychological constructs. As the section on Reliability and Validity of Scales (*see* VALIDITY AND RELIABILITY) in this encyclopedia indicates, it is very important to develop scales that measure what they are supposed to measure. Testing the reliability and validity of scales is a time-consuming task. Before writing original questions for every issue

Annexure III

Please evaluate the career planning center by circling one number for each of the following items.

	Highly Disagree				*Highly Agree*
Accessible	1	2	3	4	5
Up-to-date job information	1	2	3	4	5
Helpful staff	1	2	3	4	5
Useful in career planning	1	2	3	4	5
Networking resources	1	2	3	4	5
Courteous staff	1	2	3	4	5

that needs to be covered in the questionnaire, it is advisable to look for scales that are already developed by other researchers. For example, if we are interested in measuring consumers' perception of quality of service provided, there are already published scales in academic journals that are reliable and valid. A good source of such a scale is Bearden, Netemeyer, and Mobley (1999).

Asking questions about demographics. Sometimes, there is a tendency to collect very detailed demographic data even if it may not be necessary for the purpose of the study. In addition, researchers may want to collect exact information (income, age, etc.) even though it may not be necessary to collect such precise information. Many respondents do not like to disclose their age, income, and other information. Asking for actual income and age is a big demotivator with respect to providing such information. However, if we ask the respondents to choose an answer from a range of answers, they tend to feel more comfortable in responding to such questions. Examples of two alternative ways of asking questions about income are presented below:

What is your current total family income before taxes? $_____

Please check the category that reflects the best estimate of your current total family income before taxes?

__ Under $50 000
__ $50 000–$59 999
__ $60 000–$69 999
__ $70 000–$79 999
__ $80 000–$89 999
__ $90 000–$99 999
__ $100 000 or more

Labeling each option. Whenever possible, it is better to label each option of a rating scale rather than labeling only the end points, which is the common practice. Labels tend to

Annexure IV

How important is a "variety of choices" of fresh vegetables when deciding where to shop for vegetables?

Very important	1	2	3	4	5	Not important at all

Very important	Somewhat important	Neutral	Somewhat unimportant	Not important at all
1	2	3	4	5

provide clear guidance for the respondents when choosing a point on a scale, making it unnecessary for them to make their own assumptions about what the points represent. Examples of two scales, one with and one without all labels is presented in Annexure IV.

Availability of option to indicate uncertainty of answer ("don't know" or "no opinion"). It is quite possible that given the nature of the research topic, some respondents may not have any opinion on a subject or may not know the answer to a question. This is something that can be monitored during the pilot test of the survey (described later). On the basis of the result of the pilot test, wherever needed, the researcher should offer the option of "don't know" or "no opinion" or "undecided" so that respondents are not unnecessarily forced to choose or to give an answer.

Single versus multiple responses to a question. Sometimes, a question may require choosing only one of several options provided or choosing any combination of responses. Instructions for completing the survey should make it very clear what is expected of the respondents.

Wording the questions. The quality of information we obtain from a survey largely depends upon the quality of the questions. If the respondents do not assign the same meaning to a question that the researcher intended to assign, it does not matter whether the respondents answer that question. Although it is perhaps almost impossible to prepare a manual on how to write good questions, we can identify a number of common mistakes that are made when writing questions. They are briefly described below. The next time one sees a questionnaire, chances are that one will notice at least one of the following mistakes made in writing the questions.

Avoid leading questions. A question that also seems to suggest the answer is called a *leading question*. For example, "Do you think that outsourcing of jobs is a good idea when it leads to Americans losing their jobs?" The question is loaded with a suggested answer or tends to lead respondents to give a "right" but biased answer.

Avoid double-barrel questions. Consider this question: "How do you like the hamburger and fries at McDonald's?" This is one of the most common mistakes committed in writing questions. This questions has two questions embedded in it. There are four possible answers to this question: (i) I like both the hamburger and the fries; (ii) I dislike both of them; (iii) I like the hamburger but do not like their fries; and (iv) I dislike their hamburger but like the fries.

Exclude burdensome questions. Sometimes, a question may demand too much time and cognitive effort on the part of respondents. For example, a question such as "How much money do you spend on gasoline per year?" may result in wild estimates simply because the respondents do not know the answer or do not bother to spend the time to estimate the expenditure.

Avoid ambiguous options. Consider this question: "How often do you shop in a department store in a month?"

1. Never	2. Occasionally	3. Sometimes
4. Often	5. Regularly	

What is "occasionally" and what is "sometimes" may vary among respondents depending upon the reference point they use for comparison. These words do not refer to any specific numbers.

Define the construct clearly. What is your income? _____. The question does not specify individual or household income, before or after income tax, income from salary or total income from all sources, for which year, and so on.

Do not Assume. Sometimes, the researcher may have an assumption underlying a question, which is not stated in the question: "Are you in favor of improving the quality of food offered by this restaurant?" Here the assumption is that if the respondent is in favor of improving the quality of the food, she/he is also willing to pay a higher price, which may or may not be true. It also assumes that the quality of food needs improvement even though that may not be the case for some or for all respondents.

Avoid overlapping response options. Options should not overlap. Consider the following instruction and the options provided: "Please circle the number representing your age category."

1. Less than 20
2. 20–30
3. 30–40
4. 40–50
5. Over 50.

In this question, two people aged 30 may circle two different numbers (2 and 3).

Avoid industry/profession-specific jargon/ Acronyms. Not everybody is familiar with industry (bytes for computers) or profession-specific jargon or acronyms (e.g., GPS).

Use "*Other (Please specify)*_____" sparingly as an option. Inserting "other" as an option frequently simply indicates that the researcher has not spent time thinking about possible responses. It is better to conduct a pilot test of the questionnaire and present options that capture a large majority of the respondents for a given question. Putting "other _____" is unlikely to generate responses from most of the respondents unless an important option is missing from the list of options provided.

Options need to be balanced. Unbalanced response categories is a common problem in questionnaire design. Consider the following question and the response categories: "How satisfied are you with your current computer?"

1. Very satisfied
2. Moderately satisfied
3. Unsatisfied.

In this example, there are two "satisfied" response categories and only one "unsatisfied" category. Similarly, options such as "excellent," "good," "fair," and "poor" are also unbalanced. They need to be balanced to avoid creating response bias.

Options should be mutually exclusive. Consider the following question: "What is your marital status?"

1. Single
2. Currently married
3. Separated
4. Living with a partner.

There are two problems with the options provided. First, the option "Single" does not distinguish between those who have never married and those who are widowed or divorced. Second, the last option (Living with a partner) could possibly apply to the other three categories.

Include filter questions if necessary. Answers to filter questions can be used to eliminate those respondents who do not meet certain criteria so as to judge their answers to be meaningful. For example, one can include a question regarding the level of usage of a product. The researcher may decide to exclude the light users or analyze their responses separately since they may not know enough about the product to give meaningful answers.

Organization of questions. The organization of questions in a questionnaire plays a very important role in keeping the respondents motivated to complete the survey. A proper organization of questions is necessary not only to motivate the respondents, but it will also reduce the cognitive effort required to complete the survey. Dillman (2000) and other researchers have provided a number of suggestions to organize the questions in a survey.

- One should consider dividing a questionnaire into sections whereby each section represents the questions that are similar in content. For example, questions related to buying behavior may be presented in one section, whereas another section may deal with consumers' beliefs about different aspects of a market.
- Questions should be presented in order of importance to the respondents. Many respondents do not complete surveys after answering the first few questions. For this reason, there is a tendency on the part of survey designers to put the questions in a sequence of most important to least important questions as perceived by the

researcher, hoping that at least the most important questions will be answered. Actually this strategy is not only likely to reduce the overall response rate of the survey this is also likely to lead to many incomplete surveys. To encourage the respondent to complete the survey, the questions should be arranged in the order of importance to the respondent. Once he has invested time in responding to questions that are important to him, he is likely to answer other less-important questions as well.

- First, present questions that are somewhat easier to answer. Researchers have recommended beginning with easy questions just to make the respondents feel comfortable and become interested in the survey even though those questions may not be relevant for the main purpose of the study. Once a person invests some time responding to some questions, the likelihood of completing the survey by that respondent increases significantly over that of a person who has received the survey questionnaire but has not started responding to the questions.
- Whenever possible, questions or sections of questions should be arranged in such a way that there is some kind of natural flow. For example, if we are conducting a survey about consumers' buying behavior associated with cars, a natural flow would be to first ask questions about their "knowledge" (thinking), then their "preference" (feeling), and then their "behavior" (doing).
- Demographics/classification questions should appear at the end unless any demographic questions are required to prescreen the respondents.

Physical appearance of the survey. A number of factors should be considered regarding the physical appearance of the self-administered surveys (mail and online surveys).

Font used in the questionnaire. To cram a maximum number of questions into a survey, there is a tendency to use small fonts. If the respondents have a hard time reading the questions, they will not take the extra time needed to read the questions and complete the survey questionnaire. Fonts should be large enough so that the questions can be read with ease.

Precoding of questions. Coding of answers refers to assigning numbers to each response category so that some kind of statistical analysis can be performed on the data. If computers are used to collect the data, the precoding of closed-ended questions becomes essential. For example, if we conduct a web survey or computer-assisted telephone survey, responses to each closed-ended question are precoded so that the two tasks of collecting data and generating the data file can be done simultaneously. However, for mail surveys, options can be left without any numerical coding of those options. It is recommended that, wherever possible, the numerical codes could be included in the questionnaire itself rather than assign them after collecting the responses. This will help reduce the time it takes to code the responses later.

Provide clear instructions for each question. Do not make any assumptions regarding respondent's ability to complete the survey without clear instructions. Any given questionnaire may have different kinds of questions such as open-ended versus closed-ended, rating versus ranking questions, questions requiring a single response versus multiple responses, to name a few. Each question must be accompanied by clear instructions regarding what the respondent needs to do in order to respond to that question. For example, if one does not allow a tie in a ranking question, one needs to specify that the respondent should not use the same number to rank two objects.

Pre-testing of the questionnaire. Pretesting a questionnaire is an extremely important and absolutely necessary step in designing a quality questionnaire. The main purpose of pretesting the questionnaire is to ensure that the target respondent's interpretation of each question matches the intended meaning of the question from the viewpoint of the researcher and the respondent is able and willing to answer each question. In other words, pretesting is necessary to decrease both sampling and nonsampling errors in a survey study. Pretesting is usually done on a small sample (about 30) of respondents, representative of the population of the

study. Before finalizing the survey, it is important to test the survey on several dimensions:

1. Do the target respondents understand each question? It is quite possible that there are special words or technical language used in the questionnaire which are not in the vocabulary of respondents.
2. Are there any any questions that are likely to be perceived as offensive?
3. Are the respondents able to provide the information asked in the survey (e.g., calorie intake everyday, money spent on gasoline per month, etc.)?
4. Are there certain questions that respondents tend to leave unanswered in the pilot survey? During the debriefing session with the respondents who participated in the pilot survey, ask for reasons for such behavior.
5. Are there any questions that are too long to understand?
6. Are there any questions that seem to disturb the natural flow of other questions? It is quite possible that some questions are located in inappropriate places in the questionnaire.
7. Are there questions that the respondent is hesitant to answer accurately?
8. Did the respondent understand the instructions for responding to each question?
9. Are there questions that are likely to generate socially desirable answers?

Pretesting is also important from the view point of assessing the questionnaire as a whole in terms of (i) the time it takes to complete the survey, (ii) checking to see whether additional questions are needed to cover the scope of the study, and (iii) checking the reliability and validity of new scales.

The pretesting should be done by taking a representative sample of the target population. This can be done in several ways. One way of conducting the pretest is to ask each respondent to read each question and think aloud while interpreting the meaning of the question and then answer that question. This process reveals the thought processes of the respondent and helps evaluate the quality of the questions. Another way of conducting the pretest is to involve more than one person in the administration of the survey to a small sample of respondents. In addition to completing the survey, each respondent is asked to provide comments on different aspects of the survey during a debriefing session. A debriefing session essentially provides an opportunity for the researcher to explain the purpose of the study to the participants, to understand how they feel about the survey, and to solicit comments, if any, from them. Once the surveys are completed, the survey administrators collectively discuss the opinions of and problems encountered by the respondents and correct/improve the survey accordingly. Another way of pretesting the survey is to conduct a face-to-face interview with a small sample of respondents. To do this, after the survey is completed, discuss the survey with the participating respondents in a debriefing session.

When the survey questions are fixed after pretesting, an important step in finalizing the survey is to compare the survey with (i) the information desired listed in Table 1, which shows the list of hypotheses and research questions and (ii) the information likely to be generated from the survey to test hypotheses or answer research questions. It is quite likely that after several iterations of writing, rewriting and finalizing questions, it will be discovered that the current version of the questionnaire does not obtain some of the important information needed for the study. In that case, we need to add the missing questions. At the same time, there is very high likelihood of including questions that are not necessary. Therefore, for each question, we need to ask: "Is the question absolutely necessary given the objectives of the research?" Questions that are simply guided by the motive of "it will be really interesting to know," rather than "necessary to know given the purpose of the study," should be eliminated.

Increasing the Response Rate

Response rate is defined as the number of usable surveys received divided by the number of surveys actually delivered or the number of respondents actually contacted. It is generally believed that other things remaining the same, a high response rate is better than a lower response rates in order to increase the sample size and thus

reduce the sampling error. However, it must be noted that it is more important to have a sample that is representative of the target population than a high response rate. A high response rate from a biased sample simply adds to the nonsampling error to research findings. Here are a number of things a researcher can do to increase response rate.

1. Specifying the sponsor and purpose of the study
2. Sending advance notification before actual survey
3. Calling back or sending reminder
4. Offering motivation: monetary reward, summary of results, donation to charity of respondents' selection
5. Reducing the time required to complete the survey
6. Reducing the effort required to complete the survey
7. Providing assurance of maintaining respondents' privacy and confidentiality of their responses
8. Ensuring that the promise regarding incentives are fulfilled.

In conclusion, a poorly designed questionnaire is likely to affect the both quality and quantity of information collected from the survey, ultimately increasing the total error in study findings. By following the simple guidelines presented in this article, we can improve the quality of the questionnaire significantly.

Bibliography

Bearden, W.O., Netemeyer, R.G., and Mobley, M.F. (1999) *Handbook of Marketing Scales*, 2nd edn, Sage Publication, Newbury Park.

Czaja, R. and Johnny, B. (1996) *Designing Surveys: A Guide to Decisions and Procedures*, 2nd edn, Pine Forge Press, Thousand Oaks.

De Leeuw, E.D., Hox, J.H., and Dillman, A.D. (eds) (2008) *International Handbook of Survey Methodology*, Lawrence Erlbaum Associates, New York.

Dillman, D.A. (2000) *Mail and Internet Surveys: The Tailored Design Method*, John Wiley & Sons, Inc., New York.

Krosnick, J.A. (1999) Survey research. *Annual Review of Psychology*, **50**, 537–567.

Lewis-Beck, M.S., Bryman, A., and Liao, T.F. (eds) (2004) *The Sage Encyclopedia of Social Science Research Methods*, vols 1, 2 and 3, Sage Publications, Thousand Oaks.

Malhotra, N.K. (2004) *Marketing Research: An Applied Orientation*, 2nd edn, Pearson-Prentice Hall, Upper Saddle River.

sampling techniques

Wagner A. Kamakura

It is quite rare that a marketing researcher has the luxury of studying an entire population of consumers. Most of us must resort to drawing inferences about the population from data gathered from a sample. Therefore, before gathering the data we must first select a sample of consumers that is representative of the population as a whole.

The Sampling Frame

To study a representative sample of the population, the researcher must have an accurate definition of the population from which the sample is to be drawn. While this seems rather simple and obvious, in many cases the researcher does not know precisely what the population might be. For example, if the purpose of the study is to assess the sales potential for a new energy drink, the population includes all consumers who might have a need for such a drink, making it difficult to identify members of the population a priori. Sometimes, a representative sample might come from a different population from the one the researcher wishes to infer about. For example, a wireless telecom might want to draw a sample from its current customers to learn how to acquire new customers who are likely to be loyal and valuable to the firm, thereby using a sample from a different population (current customers) to draw inferences about the focal population (nonusers or customers from competitors).

Because it is not always possible to identify every member of the population, the researcher must use a *sampling frame* from which to draw the representative sample. A sampling frame is a *subset of the population that the researcher believes to be representative of it*. Because only members of the population contained in the sampling frame will be chosen for the sample, it is critical that the sampling frame is representative of the population. This sampling frame can be a list of population members (such as a telephone book, or e-mail address list), but does not have to necessarily enumerate population members. For example, a city map might be used as a sampling frame, from which the researcher draws a sample by selecting city blocks and sending interviewers to survey every third house in the block. Conversely, a sampling frame might also contain nonmembers of the population, such as a list of telephone exchanges and area codes that might include nonresidential numbers or those not in service, which can be dropped from the final sample.

Sampling Schemes

Once the researcher establishes a sampling frame, he or she must decide how members of the population will be selected from those included in the sampling frame. There are two basic sampling schemes that are discussed in more detail elsewhere: *probability* (*see* PROBABILITY SAMPLING) and *nonprobability* (*see* NONPROBABILITY SAMPLING) sampling schemes. In the probability sampling scheme, each member of the population contained in the sampling frame has a known (but not necessarily equal) probability of being included in the sample. If the sampling frame is representative, random sampling is expected to produce a representative sample of the population. In the nonprobability sampling scheme, sample selection is not left to chance, but is based on some criteria defining whether members of the sampling frame are of interest for the study at hand.

Given that the researcher's goal is to draw inferences about the population from the sample, a major concern is the discrepancy between the measurements observed in the sample and their respective population values, which are unobservable. *Systematic variation* or *sampling bias* occurs when subjects selected for a sample over-/underrepresent the population features under study. For example, a sample of telephone interviews reaching people at home during the evening is likely to produce a biased sample for a life-style survey, because it will underrepresent consumers enjoying life away from home. Systematic variation or sampling bias may occur in different forms in a marketing research study:

- *Nonresponse bias*: In a typical survey study, low response rates may render the effective sample of respondents nonrandom, even when a random sampling scheme is used

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

to draw the original sample. Low response rates may also render the effective sample nonrepresentative, when the profile of respondents deviates from nonrespondents. This sample bias is often assessed by attempting to contact a portion of the unresponsive sample elements and comparing respondents and nonrespondents on the main variables being measured.

- *Coverage bias*: When certain types of population members do not appear in the sample frame, the observed sample values deviate from the population owing to differences between covered and noncovered units. This type of bias is a concern in certain web surveys when the target population is not fully covered by this particular medium.
- *Selection bias*: When some members of the population have a higher/lower probability of sample selection than planned by the researcher. For example, random-digit dialing is likely to produce selection bias because some households have multiple phone numbers making them more likely to be sampled than households with only one phone number.

When sampling bias cannot be avoided, data collected from the sample must be weighted to correctly represent the population. For example, a sample collected by intercepting shoppers at a shopping center is likely to overrepresent frequent and slow (those who linger longer) visitors to the center. The researcher may reduce this selection bias by weighing each sampled respondent inversely to the number of trips in the past six months and minutes spent at the center prior to the interview.

Random variation or error occurs due to chance, rather than human biases, and can be assessed statistically when the sample is drawn randomly (*see* UNIVARIATE TECHNIQUES). Because selection of sample elements is nonrandom, nonprobability sampling does not allow for the assessment of sampling errors.

Sampling schemes also differ on whether sample elements are collected with replacement (a member of the sampling frame may be sampled more than once within the same study) or without replacement; most marketing research studies use samples without replacement. Sample biases due to nonresponse can also be mitigated by weighting respondents in proportion to their over-/underrepresentation in the final sample relative to the population profile.

Sampling Methods

A variety of sampling methods are available to the marketing researcher, following the probability and nonprobability sampling schemes discussed above. The choice of one of these methods depends on cost considerations, the type of sampling frame used to draw the sample, the availability of additional information about each member of the sampling frame, and whether it is important to assess sampling error. For more details on sampling methods, see Govindarajulu (1999) and/or Cochran (1977).

Convenience or opportunity sampling. This is probably the most common form of sampling method used in commercial marketing research. It involves sampling consumers who are members of the target population and happen to be more easily available to the researcher. The main reason for the popularity of this nonprobability sampling method is cost. Another common justification for using convenience samples is that response rates to marketing research are often low, thereby making the effective samples obtained from probability sampling methods (discussed below) as representative/non-representative as those obtained with this simple and less expensive method.

One form of convenience sampling particularly useful to sample hard-to-find populations is *snowball sampling*, where sampled subjects are asked to help recruit additional members to the sample. Another form of convenience sampling is *judgment sampling*, in which the researcher selects the sample based on judgment. For example, a researcher may elect to run a test market in a few "representative" metropolitan areas, even though the population of consumers includes all cities.

Because bias and representativeness are major concerns for this type of sampling, the researcher must ascertain that the sample profile does not depart considerably from the population on relevant aspects to the study at hand.

Simple random sampling. In a simple random sample, members of the sampling frame are randomly drawn for the sample with the same probability, in an attempt to minimize sampling bias and obtain a representative sample (*see* PROBABILITY SAMPLING). However, simple random drawing may result in the over- or underrepresentation of small segments of the population. For example, suppose that ethnicity is a critical aspect in a particular study and Asians represent 2% of the population. A sample of 100 members randomly drawn from the population may contain no Asian or 5 Asians, leading to no representation or overrepresentation of this particular segment. Simple random sampling might also be impractical in some cases, because it requires that the researcher assigns equal selection probability to each member of the sampling frame.

Systematic random sampling. A more convenient way of drawing a random sample from a sampling frame is by systematic random sampling. Suppose the sampling frame is a list containing 6000 members and the researcher needs a random sample of 200. The first step in systematic random sampling is to define the sampling interval, which, in this example, is 30 (6000/200). The second step is to draw a number randomly from the sampling interval (between 1 and 30 in this example), which defines the first sample element. For example, suppose the random number was 16. Then, the 16th member listed in the sampling frame will be the first selected member for the random sample. From this random starting point, the researcher selects every 30th member listed in the sampling frame, until reaching the end of the list. Because the first member was chosen at random, the resulting sample is a simple random sample.

If the sampling frame is sorted according to some consumer characteristic of interest (e.g., ethnicity), the systematic random sampling procedure described above will ensure that each segment (e.g., Asians, Hispanics, etc.) is represented in the same proportion in the sample as it is in the sampling frame, leading to a proportionate stratified random sample, as described next.

Stratified random sampling. In cases where the population is categorized into relevant segments or *strata* and the sampling frame contains information about this segmentation, a stratified random sample can be drawn to ensure that each *stratum* or segment is represented in the same proportion within the sample as in the sampling frame. The first step is to segregate the sampling frame into strata of interest (e.g., ethnic groups). Each stratum is then treated as an independent subpopulation, and a simple random sample drawn from each. Aside from ensuring that each stratum is properly represented in the sample, this stratified sampling produces more efficient statistical estimates, if each stratum is more homogeneous relative to the differences across strata, in terms of the construct being measured.

Since each stratum is treated as an independent subpopulation, different random sampling methods may be applied to each. A common variation of stratified random sampling is *nonproportionate stratified sampling*, where the relative sizes of the strata in the sample depend on their heterogeneity, so that more consumers are sampled within heterogeneous strata than in more homogeneous ones, to improve overall statistical efficiency.

Another form of nonproportionate stratified random sampling is *choice-based sampling*, where the sample is stratified on the choice outcome and a sample is taken from each strata so that the less frequently observed strata is overrepresented in the sample. A statistical model is then built on this biased sample, and the effects of predictor variables on choice can be estimated with more precision with this choice-based sample, compared to a random sample. This sample design is often used to study rare events, and the final results are adjusted to correct for the oversampling of the rare events.

Cluster random sampling. When the sampling frame is naturally organized into "minipopulations," each with a profile that represents the population, a cost-effective random sample can be drawn using cluster random sampling. For example, a textbook publisher interested in interviewing students of secondary schools may consider each secondary school as a "minipopulation" and draw a cluster sample in two

or more stages. The first stage would be to draw a random sample of secondary schools from a sampling frame. In a second stage, the researcher could draw a sample of classrooms in each of the sampled schools. A final sample would then be collected in a third stage, by randomly drawing a sample of students in each of the sampled classrooms. Cluster random sampling typically results in lower costs, but can increase the sampling error relative to simple random sampling, depending on how the clusters differ from each other, compared to within-cluster variation. In contrast to stratified random sampling, in cluster sampling one wants high within-cluster variance, relative to between-cluster variance. In other words, in cluster sampling one wants each cluster to be as representative of the population as possible. Cluster sampling is used to draw random samples for door-to-door personal interviews by first drawing a random sample of city blocks and then drawing random samples of houses within each of the sampled blocks. Another example of cluster sampling can be found in random-digit dialsing for phone surveys, where one first draws a random sample of exchange/area-code combinations, and then draws random suffixes within each of the chosen exchange/area-code combinations.

Quota sampling. This nonprobability sampling method is similar to stratified random sampling. The first step, of classifying the sampling frame into strata, is the same, but instead of drawing members from each stratum or segment at random, judgment is used for this selection in quota sampling. This type of sampling is common in intercept surveys, where the interviewer is placed at some location and asked to sample a given quota of consumers fitting a specific profile (e.g., 200 female consumers between 20 and 30 years of age). Because the selection of respondents fitting the specified profile is left to the interviewer, the sample may be biased by the interviewer's decision process; for example, the interviewer may choose to interview those who appear more approachable or friendly.

Once the sampling frame has been defined and the sampling method chosen, the next step is to decide for the sample size, which is the subject of another article (*see* STATISTICAL APPROACHES TO DETERMINING SAMPLE SIZES) .

Bibliography

Govindarajulu, Z. (1999) *Elements of Sampling Theory and Methods*, Prentice-Hall, Upper Saddle River, NJ.

Cochran, W.G. (1977) *Sampling Techniques*, 3rd edn, John Wiley & Sons, Inc., New York.

statistical approaches to determining sample sizes

Wagner A. Kamakura

The main reason for a researcher to draw a random sample is to make statistical inferences about the population, on the basis of a more affordable and more manageable set of observations than canvassing the entire population (*see* SAMPLING TECHNIQUES). The main question then is how large the sample should be to allow for this inference regarding the population. All else being equal, a larger sample will lead to more precise estimates about the population. Therefore, before determining the sample size, the researcher must first decide how much precision is required in the inference about the population.

Sample Size to Estimate the Population Mean or Proportion

If the purpose of the study is to estimate the population mean (or proportion) on a continuous (or binary) variable (*see* UNIVARIATE TECHNIQUES), the researcher must specify the maximum accepted error and the desired confidence level. In other words, the researcher wants the population mean (or proportion) to be within *ERROR* units from the sample result, with a confidence level *CL* (typically expressed in a proportion).

If a large enough sample is collected, the distribution of the sample mean $\overline{x}$ (or proportion p) approaches a normal distribution centered around the unknown population mean (or proportion) and a variance equal to the population variance σ^2 divided by the sample size n. In other terms, the difference between the sample and population means (or proportions), which we defined as *ERROR*, is normally distributed with zero mean and a standard deviation equal to $\sigma_{\overline{x}} = \sigma/\sqrt{n}$ (or $\sigma_p = \sigma/\sqrt{n}$), where σ^2 is the population variance and n is the sample size.

Therefore, the researcher wants a sample size such that the probability of the population estimate falling within *ERROR* units from the sample result is *CL*, or

$$\Pr\{\overline{x} - ERROR \leq \mu \leq \overline{x} + ERROR\} = CL = \Phi\left\{\frac{ERROR\sqrt{n}}{\sigma}\right\} \quad (1)$$

$$\Pr\{p - ERROR \leq \pi \leq p + ERROR\} = CL = \Phi\left\{\frac{ERROR\sqrt{n}}{\sigma}\right\} \quad (2)$$

Given that our goal is to determine the sample size needed to achieve a precision of *ERROR* units at a confidence level of *CL*, we solve for N, obtaining

$$n = \frac{z^2\sigma^2}{ERROR^2} \quad (3)$$

where

$$z = \Phi^{-1}\{CL\} \quad (4)$$

A commonly used confidence level is 0.95 (leading to the 95% confidence interval), with a corresponding $z = 1.96$, which is often rounded up to 2.

The sample-size formula above is inconvenient because it uses knowledge about the population variance before the sample is collected. If the sample is being collected to estimate a population proportion, the largest possible population variance is 0.25 and therefore, the largest needed sample size would be $n = \frac{(0.25)z^2}{ERROR^2}$, which for $CL = 0.95$ simplifies to $n = \frac{1}{ERROR^2}$, a simple formula commonly used to define the sample size for opinion polls and other surveys. For example, if the survey is to produce estimates with a margin of error of 5% and confidence level of 95%, the required sample would be 400 (1/0.0025).

If the purpose of the sample is to infer the population mean, a rough estimate of the population variance can be obtained by approximating the standard deviation σ as one-sixth of the likely range of the focal variable in the population. A better approach is to collect the sample sequentially, starting with a reasonably small sample, using the variance in this sample as a substitute for the population variance in the sample-size formula, drawing more members to the sample, and repeating the process until the final required sample is collected.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Sample Size for Stratified Sampling

Proportionate stratified sampling. To draw a proportionate stratified sample, the required sample size for the smallest strata in the population is determined. For example, if the sample is being stratified by marital status and there are fewer widows in the population, the researcher starts by setting the sample size for widows. The standard formulas for simple random sampling are used to calculate the sample size in the smallest stratum for the desired error level and level of confidence for this stratum. Next, the total sample size is calculated by dividing the sample size of the smallest stratum by its proportion in the population. For example, if it is concluded that the researcher needs to sample 100 widows and they represent 5% of the population, then the total sample size is 2000 (100/0.05). Finally, the sample size is calculated in each of the other strata needed to achieve the same ratio in the sample as in the population.

Nonproportionate stratified sampling. In this case, the researcher wants to find the relative sizes of each stratum that provide the most precision, given a fixed sample size. The solution to this problem is the optimal allocation rule where the relative size of each strata in the sample is directly proportional to its size and variance in the population:

$$n_i = \frac{n(N_i\sigma_i)}{\sum_{i'} N_{i;}\sigma_{i'}} \tag{5}$$

where n is the total sample size; N_i is the size of stratum i in the population; σ_i is the population standard deviation within stratum i.

Sample Size for Hypothesis Tests

A common problem facing marketing researchers is to draw a sample to perform a hypothesis test (*see* HYPOTHESIS TESTING RELATED TO DIFFERENCES – PARAMETRIC TESTS) with a known statistical power, given a predetermined Type I error α. For example, the researcher wants to collect data on a continuous variable from a random sample to test the following null hypothesis about the population

$$H0 : \mu = \mu 0 \tag{6}$$

against an alternative hypothesis

$$Ha : \mu = \mu_{\mathrm{a}} \tag{7}$$

The researcher wishes to reject the null hypothesis H_0 with a probability of at least $1 - \beta$ when the alternative H_{a} is true. The probability $(1 - \beta)$ is the power of the test, or the ability to reject the null hypothesis when it is false. However, the researcher also wants to control the error in the test, so that it rejects H_0 when it is true with a predetermined probability α (known as the *Type I error*).

Given the desired Type I error α, the researcher's decision rule is to reject the null hypothesis H_o when $\overline{x} \geqslant z_\alpha \frac{\sigma}{\sqrt{n}} + \mu_o$, where $z_\alpha = \Phi^{-1}\{1 - \alpha\}$ and $\Phi^{-1}\{\}$ is the inverse standard normal cumulative density function.

However, the researcher also wants a power of $1 - \beta$, thereby requiring that

$$\overline{x} \geqslant z_\beta \frac{\sigma}{\sqrt{n}} + \mu_a \tag{8}$$

where

$$z_\beta = \Phi^{-1}\{1 - \beta\} \tag{9}$$

Combining the Type I and power conditions, one obtains the same size

$$n = \sigma^2 \left(\frac{z_\alpha + z_\beta}{\mu_a - \mu_o} \right)^2 \tag{10}$$

Bibliography

Sudman, S. (1976) *Applied Sampling*, Academic Press, New York.

Cochran, W.G. (1977) *Sampling Techniques*, 3rd edn, John Wiley & Sons, Inc., New York.

survey research

Jagdish Agrawal

Survey research refers to the method of collecting data by means of a set of questions commonly known as a *questionnaire*. The questionnaire can be administered through personal interviews (face-to-face interviews) or through mail, telephone, fax, or *via* online (electronic) surveys. Although there are other well-known methods of collecting data (such as experiment and observation), conducting surveys by means of a questionnaire is perhaps the most popular means of collecting data for academic, industry (private), and government research. The availability of the Internet as a new mode of surveying has further made the survey method of collecting data more popular than ever before. This section describes the four most popular methods of conducting surveys: personal (face-to-face) including mall-intercept interview, telephone interview, mail survey, and online (electronic) survey. Table 1 presents a comparative evaluation of these four methods.

Personal (Face-to-Face) Interview

Collecting data through a personal interview is the oldest method of conducting a survey. The US government still relies on personal interviews for collecting census data every 10 years. The personal interview involves face-to-face interaction between an interviewer and an interviewee. The interview may involve a single respondent or multiple respondents. For example, an interviewer may be interested in knowing the roles different people in a household play while buying a new house or a new car. The interview may take place at the home of the interviewee. Sometimes, the interviewer may simply intercept potential interviewees in shopping centers and conduct the interviews on the spot or request that they go to a research facility designed for such interviews. This practice is commonly known as *mall-intercept interviews*. Although personal interviews are usually conducted using the paper-and-pencil method of asking questions and recording responses, the computer-assisted personal interview (CAPI) method is gaining popularity. In the CAPI method, the respondent views the questions on a computer terminal and types the responses. Data collected through personal interviews are prone to data-coding and typing errors. An added advantage of computer-assisted interviewing is that software can perform internal checks for a range of valid answers, choice of a single versus multiple responses, skipping of certain questions depending upon the answer to previous questions, and so on. The system also enables the simultaneous creation of a data file of responses while interviews are being conducted.

In many countries, where the postal system is not very reliable and where a modern means of communication (such as the telephone and Internet) are not yet widely available, the personal interview is still the dominant means of collecting survey data. Even if other means of conducting surveys were as widely available as in the United States, there are times when the personal interview is still the most suitable means of collecting data. For example, the personal interview becomes somewhat of a necessity if the researcher wants to show an object (such as a product or advertisement or any other promotional material) and get the respondent's reaction to specific questions regarding such objects. For this reason, mall-intercept interviews are very popular. Personal interviews are also suitable for open-ended questions that may require lengthy answers and/or require further probing of answers. Once a rapport is established between the interviewer and the respondent, the length of the survey is less problematic in the personal interview than in other methods. Because of the presence of the interviewer, personal interviews also allow the flexibility of asking questions in different formats. The interviewer can help clarify the questions and response options.

In spite of a number of distinguishing advantages of personal interviews over other methods of collecting data, the personal interview is the costliest method of collecting survey data. It requires hiring well-trained interviewers. Not only do the interviewers have to persuade some people to participate in the survey, they also have to administer the survey and record responses without any personal bias. Whether done at a residence or at shopping mall, getting the cooperation of people to participate in a personal interview is a daunting task. Unlike

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Comparisons of methods of conducting surveys on a few major dimensions.

Characteristics	*Mail*	*Face-to-Face*	*Web-Based*	*Telephone*
Response time	Long	Longest	Shortest	Short
Costs	Low	Highest	Lowest	Moderate
Data processing required	Yes	Yes	No	Yes
Ability to show stimuli	Limited	Very high	Moderate	None
Length of survey	Long	Longest	Short	Short
Ability to handle complex questions	Limited	Highest	High	Lowest
Response rate	Moderate	High	Low	Moderate
Coverage of general population	Very high	High	Limited	Very high
Ability to handle sensitive information	Moderate	Limited	Limited	Moderate
Availability of data in real time	No	No	Yes	No
Easy to follow up	Easy	Most difficult	Easiest	Easy
Ease of doing it operationally	Moderate	Difficult	Easy	Easiest
Ability to probe responses	Limited	High	Limited	High
Use of open-ended questions	Limited	Ideal	Limited	Moderate

The evaluation of methods is based on the assumption that data collection is done in-house rather than by outside vendors.

other methods, including the telephone survey, where the respondents do not see the surveyors, face-to-face interviews are also susceptible to influence due to the way the interviewer asks the questions, his/her personality, and demographic characteristics. Interviewers might also introduce biases in recording responses. For example, the interviewer could record the responses to open-ended questions as per his/her interpretation of the responses given by the interviewee. One of the common problems in the face-to-face interview is that many interviewers tend to use their own language when asking questions rather than adhering to the language specified in the survey questionnaire. Obviously, conducting personal interviews requires the need for very well-trained interviewers who can not only persuade respondents to participate in interviews but also ask questions and record answers in a very objective manner.

Shopping malls are convenient locations to find people to conduct face-to-face interviews with. The mall-intercept personal interview essentially involves intercepting someone in a shopping mall to conduct a personal interview. The interview may take place right where the person is intercepted or in a research facility located inside the mall. Those who participate are rewarded monetarily or with coupons. Apart from being able to locate people for interviews, mall-intercept interviews provide a very good opportunity to collect data on concept and taste tests. Unlike the face-to-face interview conducted at someone's residence, mall intercepts also allow for conducting marketing experiments at research facilities inside malls.

There are a number of issues related to sampling in mall-intercept surveys. First, not everybody in the target population is likely to visit malls or do shopping at malls. In addition, there is also a tendency on the part of solicitors or interviewers to approach only those who they think are likely to agree to be interviewed. This obviously creates further bias in sampling. To make sampling somewhat meaningful, it is suggested that interviews be conducted at different malls and on different days and times. It is also important to randomize the sample selection by approaching every *n*th person passing by a particular location rather than using judgments of solicitors/interviewers. The researcher can also use quota sampling to ensure that respondents of different backgrounds are represented in the sample so as to reflect population characteristics.

A low response rate compared to face-to-face interviews at residences is believed to be another challenge related to mall-intercept interviews.

Research by Hornik and Ellis (1988) shows that solicitors could use a "gaze and touch" technique to generate compliance when intercepting individuals and requesting an interview time. Furthermore, female solicitors employing this tactic tend to further increase the response rate.

Telephone Survey

Conducting surveys on the telephone is one of the most popular methods of conducting surveys in the United States. Most of us, no doubt, have been interrupted by a phone call during dinner, not realizing that the person at the other end is interested in asking a few questions about some topic that may or may not be interesting to us. This method is becoming even more popular in other countries where per capita landlines are very limited, but owing to the very low cost of using cellular phones, this is the fastest way of reaching a large number of people. Telephone surveys can be conducted with or without the help of a computer. In the traditional method, the interviewer asks the questions and records the responses on the questionnaire. Once the surveys are completed, the responses are entered into the computer for analysis. However, computer-aided telephonic interview (CATI) allows the interviewer to record the responses directly into the computer, thus reducing the time required for data coding and creation of a database. Not only does CATI offer this significant advantage over the traditional method, the computer can also direct the interviewer to the appropriate follow-up question depending upon the responses to previous questions. With or without computer assistance, the telephone survey allows interaction between the interviewer and interviewee so that the interviewer can clarify a question and address any concerns. However, since the interaction is indirect, unlike the face-to-face interview, the problems associated with interviewer influence are less pronounced in the telephone interview.

The only things required to conduct a telephone survey are a list of telephone numbers and a questionnaire with a list of questions to be asked. Perhaps this is the easiest way of reaching respondents. The interviewer can use telephone directories to obtain phone numbers or purchase a list of telephone numbers of households that meet certain demographic or psychographic criteria. Because many households have an unlisted number, one way of reaching households with and without listed numbers is to engage in random digit dialing. After selecting the area codes and exchange codes of the target population, the last four digits of the telephone numbers can be randomly generated with the help of the computer. Although it solves the problem of not reaching unlisted numbers, this method also generates numbers that are not assigned or those that belong to a nontarget group (businesses, universities, or government offices rather than residences). The response time to the telephone survey is also minimal compared to the mail survey. By using many people to conduct the surveys, it is easy to collect data from several hundred people within a few days or a week at a much lower cost than mail surveys. This benefit, of course, comes at a cost. Telephone surveys are not suitable for lengthy surveys. Because of the misuse of this method (especially by telemarketers who try to promote products under the pretext of conducting surveys), people are becoming less receptive to responding to phone calls asking to conduct a survey. Whether the telephone survey takes place after a cold call or during a prearranged time, it is important to keep the survey to a reasonable length. Telephone surveys are also suitable for simple-design survey questions. Since the respondent has to remember the question, short questions and fewer options are essential for collecting quality data using the telephone survey.

The advent of cellular phones has provided new opportunities for collecting data through telephone surveys, especially in those countries where access to telephones is very limited. Owing to technological advances linking the Internet and cellular phones via Voice over Internet Protocol, the cellular phone could become a completely unique medium of conducting surveys.

Although the random sampling method allows the researcher to generate estimates that are representative of the population within a certain confidence interval, there is no guarantee that everybody selected or approached using the random sampling method will cooperate by

responding to the survey. Depending upon the method of survey used, the nonresponse rate may even exceed 50% of the original sample. It is reported in the literature that response rate, in general, and for telephone surveys, in particular, are declining. The low response rate could be due to both inability to contact individuals by phone (not at home, wrong number, disconnected number, refusal to answer the phone) and refusals to participate in the survey by those who are contacted.

If the response rate is very low, it may introduce the error due to nonresponse bias because of the difference between the responses of those who responded to the survey and those who were not contacted or refused to participate. As a result, literature on research methodology suggests increasing the response rate, which can be achieved at an additional cost. However, analysis of results of the published studies by Groves (2006) show that there is not any necessary relationship between nonresponse rate and nonresponse bias, although a high response rate can help reduce the potential risk of bias. Groves (2006, p. 669) suggests identifying and collecting data on auxiliary variables that are. "... simultaneously correlated with response propensity and the key survey variables." This will allow the need to make necessary postsurvey adjustments in responses collected through the probability sampling method. A number of actions, which do not involve a significant increase in the cost of collecting additional data, can be taken to increase response rate: Prenotification of survey, call back by redialing at different times and days, and pilot test of questions to identify the question(s) that discourage participation. Accessing participants through random digit dialing can also help ensure a representative sample to allow for postsurvey adjustments (*see* SAMPLING TECHNIQUES).

Mail (Postal) Surveys

Similar to telephone surveys, mail surveys are also very popular methods of collecting survey data. Mail surveys are the least intrusive of all the survey methods. They are completely self-administered by the respondent without any kind of assistance from the researcher. As a matter of fact, e-mail surveys are essentially electronic versions of mail surveys – of course, with some distinct features of their own. In mail surveys, printed questionnaires are mailed out to a sample of addresses of the target population. Depending upon the purpose of the study and the target population, addresses can be obtained from Associations (of growers, manufacturers, distributors, etc.) and telephone directories, or they can be purchased from research organizations (e.g., Dun & Bradstreet). Although there is no guarantee of who actually completes the survey, we can specify the target group in terms of its business characteristics such as type of industry, size in terms of annual sales and number of employees, and other features, to generate the sample of addresses according to the research need. Similarly, addresses are available that match certain demographics and psychographics. A common problem that arises from buying addresses from a third party is that many of the addresses tend to be outdated, resulting in a large portion of surveys being returned undelivered. This fact needs to be considered when deciding how many surveys to mail out.

Given the fact that mailed surveys are self-administered, it is very important to mail the questionnaire accompanied by a very well-articulated cover letter. At a minimum, a cover letter should identify the sponsor of the study, the purpose of the study, explain how the confidentiality of the responses and privacy of the respondents will be maintained, and include a request to return the completed survey by a certain date. Previous research has shown that prenotification helps to increase the response rate. To reduce the chances of recipients of a survey simply discarding it without paying any attention to it, a postcard or a letter prenotifying respondents about the forthcoming survey helps to build expectation about the survey. This increases the likelihood of the recipients of the survey paying at least some attention to it before deciding whether to complete the survey.

In addition to the cover letter, the surveyor needs to ensure that the respondent does not have to bear any cost of returning the completed survey. The surveyor may also choose to include an incentive, if any (cash, coupons, etc.), or a promise of incentive (e.g., donations to a charity for every completed survey). A promise to send a short summary of the research findings is

a good incentive in the case of respondents from businesses. Research on the response rate of mail surveys shows that multiple contacts made to encourage respondents to complete a survey, salience of topics covered in the survey, incentives provided to motivate response, and organizations that sponsor research have an influence on the response rate (Dillman, 2000).

Because mail surveys are self-administered, designing a mail survey is a challenge. Once the questionnaire is mailed out, the researcher has very little opportunity to interact with the respondents. Usually, the cover letter contains an 800 number so that respondent can call if she/he has any questions or concerns regarding the survey. In general, everything in the questionnaire has to be self-explanatory so that it is easy to understand each question and the response categories. In addition, it has been reported that not only the words used in questions but also the visual design of the questions have a significant impact on the response behavior in self-administered questionnaires (Christian and Dillman, 2004). The visual design may relate to the listing of items from which to choose, the labeling of all response categories on multiple-point scales, and so on. However, unlike a telephone survey, the mail survey allows the incorporation of complex questions such as the ranking of objects or the distribution of fixed points across alternatives.

Since the respondent has complete control of the questionnaire and does not have to complete the survey in one sitting, the mail survey is perhaps the best suited for addressing multiple research questions in one study. In interviewer-administered surveys (such as the face-to-face and telephone interview), surveys are completed during the first interaction. Similarly, in the case of the online survey, even though the respondent can save a partially completed survey, she/he has to sit in front of a computer to complete the survey. Therefore, it does not give the same freedom to the respondent as does a mail survey. Because it allows the respondent to take time to complete the survey, other things remaining the same, the quality of data collected through a mail survey is likely to be better. Since the respondent can complete the survey in privacy, the mail survey is also perhaps the most suitable for research on sensitive issues.

There are a large number of studies that have examined the impact of monetary and nonmonetary incentives on the mail-survey response rate. A number of researchers have also performed integrative reviews of these studies to derive conclusions that are generalizable across studies (Groves, 2006; Groves and Peytcheva, 2008). Some of the important findings of the studies dealing with this issue, and having very practical implications, are as follows:

1. Response rates to mail surveys in general can be increased through incentives.
2. Monetary incentives tend to produce a larger effect than nonmonetary incentives.
3. The nature of the relationship between the amount of monetary incentives and the response rate is not well established. Results are mixed regarding the relationship being linear or characterized by diminishing returns.
4. It is better to include incentives with the initial mailing of the survey rather than making a promise of rewards contingent upon return of the survey.

In summary, there is no doubt that incentives work to increase response rate, and a prepaid cash reward included in the first mailing of surveys produces the best results.

A natural question is: Do the incentives create any kind of response bias? Is it likely that respondents who are attracted to respond owing to incentives respond differently than those who would respond without any incentives? Although this issue has not been examined as extensively as the impact on response rate, the findings of a limited number of studies addressing this issue are encouraging. Results suggest that there is no evidence of any systematic bias in responses as a result of incentives to encourage higher response rates (Mizes, Fleece, and Roos, 1984).

Online (Electronic) Surveys

The worldwide web or the Internet is the latest addition to the types of different modes for conducting surveys. Although the coverage of the general population by the Internet is yet to

match that of the telephone and regular mail, the number of people using the Internet for one reason or another has been rapidly increasing. It is already a very effective means of surveying special interest groups of people living anywhere in the world.

There are two major modes of doing surveys online. The first is the e-mail survey, where either the questions are embedded in the e-mail itself or a separate file containing the questionnaire is attached to the e-mail. E-mails are sent out with or without seeking prior approval of the prospective respondents. The second mode is the web-based survey in which a list of respondents are sent the link to the survey via an e-mail requesting that they participate in the survey by clicking on the web address. This is the easiest way, for example, of conducting a survey of employees who may be located worldwide. E-mail addresses of different target groups of people are available from commercial companies. The survey may reside in the server operated by the researcher or by someone else (such as commercial companies, e.g., SurveyMonkey.com, Zoomerang.com). These companies, and many others, not only provide software to create/design surveys but also provide support services to take care of the complete task of the administration of the surveys, data storage, and generation of basic statistics from the responses. The survey link can also be posted on a private web site where visitors are encouraged to participate in the survey. In this section, the focus is on the web survey rather than the e-mail survey. The web survey is not simply an alternate way of delivering and receiving a mail survey. Because of the technology employed in designing and transferring surveys, it offers many distinct advantages over other methods. Depending upon the target group of respondents, the web survey is perhaps the fastest way of getting survey responses back at a very low cost.

Even though it is similar to the mail survey in that the web survey is self-administered, Internet technology allows the incorporation of certain design features in the questionnaire that are not feasible in other modes. For example, software used to design online surveys allows automatic skipping of questions depending upon the response chosen to previous questions or does not allow the skipping of a question. It can also restrict responses to valid categories (allowing single vs. multiple responses, not allowing ties in ranking, etc.). Because of these features, the researcher does not have to spend additional time checking the data for errors. Randomizing choices to reduce order bias can also be incorporated into the survey. The drop-down menu feature allows listing a large number of options without giving the impression of a lengthy survey. Online surveys also incorporate images, color, animation, and sound. These are just some of the features that make online surveys a very attractive mode of conducting surveys. There are a large number of online companies that provide survey design services and will conduct the survey and report results of the survey at a very low cost. It is not necessary to be technologically savvy to design a simple online survey.

Another great advantage of conducting a survey online is that the results can be obtained in real time. As the respondents complete the survey, responses are recorded directly into the database, ready for analysis. The researcher still has to decide how to handle responses to open-ended questions; but otherwise, this mode reduces a great deal of effort that is required to code responses and create data files. This advantage comes with few additional costs. As a matter of fact, the online survey method is the least expensive method of surveying respondents worldwide. Analysis of findings of previous research shows that the response rate of online surveys is influenced very significantly by the salience of the subject matter. Interestingly, survey length does not seem to have a significant negative impact (Cook, Heath, and Thompson, 2000).

Issues of privacy and confidentiality are the major concerns associated with online surveys – much more so than the other methods. The fact that someone unfamiliar obtained your e-mail address and sent you an unsolicited request to complete a survey is enough to raise concerns. Unlike other methods, it is very easy and costless to send multiple reminders to complete the survey. Sometimes, reminders are sent even to those who have already completed the survey. In addition, there is also the fear of the respondent's identity being revealed or not kept confidential. Researchers in this area have provided a number

of guidelines to address this issue (Andrews, Nonnecke, and Preece, 2003). For example, it is suggested that the cover letter invitation to complete the survey should be sent separately with the option to "opt in" to the actual survey, and that there should be a guarantee of protecting personal data. This process not only respects a respondent's privacy but also helps to increase the response rate.

Conclusion

As indicated by Table 1, which compares different methods on a number of dimensions, each method has its own pros and cons. Each method is more or less suitable for certain tasks. As a result, these different methods of conducting surveys are not substitutes for one another, but they are at best complementary. Some researchers have suggested using multiple methods to collect data; although using multiple methods is not always practical. What it essentially means is that we have to select the survey method very carefully based on the purpose of the study and the relative importance of the advantages and disadvantages of different methods. No matter what method is chosen to conduct the survey, it is very important to respect the privacy of respondents and honor the confidentiality of their responses.

Bibliography

Andrews, D., Nonnecke, B., and Preece, J. (2003) Electronic survey methodology: a case study in reaching hard-to-involve internet users. *International Journal of Human-Computer Interaction*, **16** (2), 185–210.

Bourque, L.B. and Fielder, E.P. (2002a) *How to Conduct Self-administered and Mail Surveys*, vol. 3, Sage Publications, Thousand Oaks.

Christian, L.M. and Dillman, D.A. (2004) The influence of graphical and symbolic language manipulations on responses to self-administered questions. *Public Opinion Quarterly*, **68** (1), 57–80.

Church, A.H. (1993) Estimating the effect of incentives on mail survey response rates: a meta analysis. *Public Opinion Quarterly*, **57**, 62–79.

Cook, C., Heath, F., and Thompson, R.L. (2000) A meta-analysis of response rates in web- or internet-based surveys. *Educational and Psychological Measurement*, **60** (6), 821–836.

Couper, M.P. (2000) Web surveys: a review of issues and approaches. *Public Opinion Quarterly*, **64**, 464–494.

De Leeuw, E.D., Hox, J.J., and Dillman, D.A. (2008) *International Handbook of Survey Methodology*, European Association of Methodology, Lawrence Erlbaum Associates, New York.

Dillman, D. (2000) *Mail and Web-based Survey: The Tailored Design Method*, 2nd edn, John Wiley & Sons, Inc., New York.

Groves, R.M. (2006) Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, **70** (5), 646–675.

Groves, R.M. and Peytcheva, E. (2008) The impact of nonresponse rates on nonresponse bias. *Public Opinion Quarterly*, **72** (2), 167–189.

Hornik, J. and Ellis, S. (1988) Strategies to secure compliance for a mall intercept interview. *Public Opinion Quarterly*, **52** (4), 539–551.

Krosnick, J.A. (1999) Survey research. *Annual Reviews: Psychology*, **50** (1), 537–567.

Malhotra, N.K. (2004) *Marketing Research: An Applied Orientation*, 4th edn, Person-Prentice Hall, Upper Saddle River.

Mizes, J.S., Fleece, E.L., and Roos, C. (1984) Incentives for increasing return rates: magnitude levels, response bias, and format. *Public Opinion Quarterly*, **48**, 794–800.

Nederhof, A.J. (1985) Methods of coping with social desirability bias: a review. *European Journal of Social Psychology*, **15**, 263–280.

analysis of variance and covariance

Sangkil Moon

Laura, a marketing researcher, is interested in examining how different MP3 players influence consumers' satisfaction in music listening. Her primary interest lies in whether there is any difference in music-listeners' satisfaction when they use different MP3 player brands. She chooses three well-known top brands to be compared (categorical independent variable with three categories) – Apple's iPod, SanDisk, and Microsoft's Zune. She also invites 300 typical consumers (overall sample size) of various demographic characteristics chosen at random (random sampling). In an artificial laboratory setting (*see* EXPERIMENTAL DESIGN), she randomly assigns each of the 300 consumers to one of three groups consisting of 100 subjects (category sample size). Now, she gives each consumer group one of the three MP3 player brands. Once she exposes the subjects to the same music for the next 30 minutes, she asks them to evaluate their satisfaction about the quality of the music they just listened to on a 10-point scale (metric-dependent variable). When Laura conducts this experiment and collects data, she can determine whether MP3 player brands influence music listening quality by using a statistical methodology named *analysis of variance* (ANOVA).

Similarly, another marketing researcher, James, exposes three groups of subjects to three distinct advertising messages (categorical independent variable with three categories) and, subsequently, asks them to rate the appeal of the advertisements on a 7-point scale (metric-dependent variable). James randomly assigns the subjects to one of the three groups, and obtains three sample means to compare. To analyze these data, he also uses ANOVA to determine whether the entire set of sample means suggests that the samples were drawn from the same general population (Hair *et al.*, 1995).

Thus, ANOVA is used to determine the probability that differences in means across multiple categories are due solely to sampling error. Quite often, data for ANOVA are effectively collected in artificial laboratory settings through careful and systematic experimental design as implied in the two examples above. This article is aimed at providing some basics of ANOVA for MBA and undergraduate students, particularly those interested in marketing research.

This article is organized in the following order. First, ANOVA is introduced with the experimental design. Then, several types of ANOVA are discussed one by one such as one-way ANOVA, *N*-way ANOVA, analysis of covariance (ANCOVA), and multivariate analysis of variance (MANOVA). Then, ANOVA is compared with MULTIPLE REGRESSION. The last section introduces some ANOVA examples applied in marketing research.

Analysis of Variance (ANOVA) and Experimental Design

ANOVA is used for determining the mean value differences of the dependent variable, (i.e., criterion variable) associated with the effects of the controlled independent variables (i.e., factors, predictor variables) after accounting for the influences of the uncontrolled independent variables (i.e., nuisance variables). In other words, ANOVA is used as a test of mean differences for two or more categories. The null hypothesis for ANOVA indicates that all category means are equal, which suggests that all the categories came from the same overall population. For example, let us suppose that the researcher is interested in examining whether heavy, medium, light, and nonusers of cereals (independent variable with four consumer groups) differ in their preference for Total cereal, measured on a 9-point Likert scale (metric-dependent variable) (Malhotra, 2007). We can use ANOVA to test the null hypothesis that the four groups of cereal consumers have the same preference for Total cereal.

One may wonder whether we can simply use the *t*-test for the same task – determining mean differences across multiple groups. Churchill and Iacobucci (2005) indicate that ANOVA is an extension of the *t*-test in two ways. First, whereas the *t*-test enables us to compare only two sample means at a time, ANOVA lets us compare the means of more than two groups simultaneously. Second, in such a two-sample *t*-test, there is only one factor that defines the group membership (e.g., buyers of brand A as

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

opposed to nonbuyers of brand A). In ANOVA, a combination of two or more factors can define the groups (e.g., buyer status and gender).

ANOVA must be composed of at least one dependent variable (e.g., preference for Total cereal) and one or more independent variables (e.g., product use (heavy, medium, light, and nonusers) and shopping frequency (frequent, regular, occasional, and never)). The dependent variables should be metric (i.e., both interval and ratio scales), whereas the independent variables are categorical or nonmetric (i.e., both nominal and ordinal scales). Such categorical independent variables are also called *factors*. A particular combination of factor levels is called a *treatment*. For example, product use is a factor and a heavy-users group in product use is a factor level. A heavy user with regular shopping frequency is a treatment as a combination of the two factors.

One-way ANOVA involves only one categorical independent variable that can influence the dependent variable. When there is only one factor involved in the analysis (e.g., shopping frequency consumer groups – frequent, regular, occasional, and never), we can examine the differences between the specified groups in the factor by one-way ANOVA. In one-way ANOVA, a treatment is considered to be equal to a factor level. If two or more factors are involved, the analysis is generalized to N-way ANOVA, where N is the number of factors involved. If the researcher needs to compare preferences for Total cereal between brand loyal and nonloyal customers in addition to among the four shopping frequency consumer groups, two-way ANOVA would be appropriate. In this two-way ANOVA analysis, there are eight different treatments from the two factors included (a combination of four factor levels from the shopping frequency factor and two factor levels from the brand loyalty factor). For example, a consumer group who are both occasional shoppers and brand loyal constitute one of the eight possible treatments.

When the analysis takes into account the consumers' interests in their body weights, measured on a 7-point scale as a *metric-independent variable*, the analysis should be upgraded to ANCOVA. In this case, the two categorical independent variables (i.e., shopping frequency and brand loyalty) are still referred to as *factors*, whereas the metric-independent variable (i.e., attitude toward nutrition) is referred to as a *covariate*. As implied in the example, ANCOVA is used to accommodate both categorical and metric-independent variables simultaneously. ANCOVA is explained in more detail later as a separate section.

From a researcher's perspective, ANOVA is particularly useful when used in combination with an experimental design (*see* EXPERIMENTAL DESIGN), where the researcher directly controls or manipulates one or more independent variables to measure their effects on the dependent variable. Without such an intended and controlled experimental design, it would be difficult to prove causal relationships between the independent variables and the dependent variable. After all, ANOVA coupled with the experimental design provides the tools necessary to determine whether observed differences between treatment groups are exactly owing to treatment effects or random sampling variabilities (Hair *et al.*, 1995). In brief, the power of the experimental design is that the researcher can actively and systematically alter the variables of interest and observe what changes take place after the manipulation (Cooper and Schindler, 2006). In contrast, to uncover causal relationships, the researcher can interview respondents after the marketing phenomenon took place, as in the ex post facto research design. Unlike the experimental design, however, this approach is required to accept the phenomenon as it is unfolded in the world.

One-Way Analysis of Variance (ANOVA)

According to the number of independent variables used in ANOVA, we can divide the analysis technique into one-way ANOVA (the simplest form) and N-way ANOVA (the generalized form). As the name indicates, one-way ANOVA contains only a single independent variable, and, accordingly, cannot have any interaction terms between multiple independent variables. One-way ANOVA can be extended into N-way ANOVA, where N indicates the number of independent variables in the analysis model. An important addition to N-way ANOVA from

one-way ANOVA is the presence of interaction terms of multiple independent variables, in most cases, interactions between two independent variables. One-way ANOVA is explained in this section and the discussion of N-way ANOVA as the general ANOVA model follows in the subsequent section.

Malhotra (2007) effectively illustrates a five-step procedure for implementing one-way ANOVA: (i) identify the dependent and independent variable, (ii) decompose the total variation, (iii) measure the effects, (iv) test the significance, and (v) interpret the results. Let us take a close look at each step one by one.

First, the researcher needs to identify the dependent variable (Y) and the only independent variable (X). The independent variable should be a categorical, nonmetric variable with c categories. Second, the researcher needs to decompose the total variation according to the variation source. In examining the category mean differences of the dependent variable one-way ANOVA divides the total observed variation (SS_y) in the dependent variable into two different components – the variation in the dependent variable related to the variation in the means of the categories of independent variable ($SS_{\text{between}} = SS_x$) and the variation in the dependent variable owing to the variation within each of the categories of the independent variable ($SS_{\text{within}} = SS_{\text{error}}$), which is expressed as follows.

$$SS_y = SS_{\text{between}} + SS_{\text{within}} \tag{1}$$

Whereas SS_{between} represents the variation between the categories of the independent variable, SS_{within} represents the variation not accounted for by the independent variable. Intuitively, we can comprehend the logic of decomposing the total variation of the dependent variable into these two subcomponents in order to examine differences in group means. In ANOVA, there are several different groups (e.g., heavy, medium, light, and nonusers). Using ANOVA, one can determine how much the sample means should vary because of random sampling variation alone. If the observed variation in the sample means is more than what would be expected by sampling variation, we can conclude that this additional variability stems from group mean differences.

Third, the researcher needs to measure the effects of the independent variable on the dependent variable using $\eta^2 = SS_x/SS_y$. That is, η^2 is a measure of the variation in the dependent variable explained by the independent variable. This measure ranges from 0 to 1. When there is no effect of the independent variable on the dependent variable, $\eta^2 = 0$. Fourth, the researcher can not only measure the effects of the independent variable on the dependent variable using η^2, but also test for its significance. In particular, in this one-way ANOVA analysis, our interest lies in testing the null hypothesis that the category means of the dependent variable are equal. This null hypothesis is tested by the following F-test statistic based on the ratio between mean square owing to X (MS_x) and mean square owing to error (MS_{error}) (Malhotra, 2007).

$$F = \frac{MS_x}{MS_{\text{error}}} = \frac{[SS_x/(c-1)]}{[SS_{\text{error}}/(N-c)]} \tag{2}$$

As noted above, the mean square is the sum of squares divided by the appropriate degrees of freedom denoted by c.

Lastly, when we interpret the ANOVA results based on the above F-test, there are two possible statistical conclusions: either the null hypothesis is tenable or rejected. If the null hypothesis of equal category means is not rejected, we conclude that the independent variable does not have a significant effect on the dependent variable. On the other hand, if the null hypothesis is rejected, we can conclude that the mean value of the dependent variable will be different across multiple categories of the independent variable. However, when the null hypothesis is rejected, the conclusion, "not all category means are equal," is not particularly explicit or useful. That is, the rejection of the null hypothesis does not automatically inform the research which means differ significantly from which other means. Therefore, when the null hypothesis is rejected, the researcher needs to learn which means are different from which other means. This search procedure is called a *multiple comparison* (*MC*) technique (Glass and Hopkins, 1984). The researcher needs to select an appropriate MC technique given the specific research

objectives. Popular MC techniques include trend analysis, planned orthogonal contrasts (POCs), the Dunnett method, the Dunn method, the Scheffé method, the Tukey method, and the Newman–Keuls (NK) method.

N-Way Analysis of Variance (ANOVA)

In addition to an interest in whether an independent variable has an effect on the dependent variable, there may be an interest in whether the factor levels associated with the independent variable are equally effective for the factor levels associated with another independent variable (Glass and Hopkins, 1984). In the context of marketing research, one is often concerned with the simultaneous effects of more than one factor. For instance, the marketing researcher is interested in whether the effects of several brand loyalty levels (e.g., high, medium, and low) vary across consumers' different income levels (e.g., high and low). If the brand loyalty level effects arising from its three different categories vary significantly between high versus low income consumers, we conclude that there is an interaction effect between the two factors of brand loyalty and income. Thus, an interaction occurs when the effects of one factor on the dependent variable depend on the category of another factor. In determining such interaction effects, N-way ANOVA can be used, where N indicates the number of factors or the independent variables. A major advantage of N-way ANOVA compared to the earlier one-way ANOVA is that it enables the researcher to examine these interaction effects between the factors.

Since the basic procedure for conducting N-way ANOVA is similar in many ways to that of one-way ANOVA shown earlier, this section highlights the main differences between the two ANOVA techniques. Although N-way ANOVA can take any number of factors N, we consider a two-way ANOVA that considers two factors (X_1 with c_1 categories and X_2 with c_2 categories) for the sake of simplicity. The total variance in this two-way ANOVA case is partitioned as follows (Malhotra, 2007):

$$SS_{\text{total}}\ (SS_y) = SS_{x1} + SS_{x2} + SS_{x1x2} + SS_{\text{within}}\ (SS_{\text{error}}) \quad (3)$$

where, SS_{x1} is SS due to X_1, SS_{x2} is SS due to X_2, and SS_{x1x2} is SS due to interaction of X_1 and X_2 as explained below.

A larger effect of X_1 is reflected in a greater mean difference in the levels of X_1, that is, a larger SS_{x1}. The same is true for the effect of X_2. By the same token, the larger the interaction effect between the two factors, the larger SS_{x1x2} is. On the other hand, if the two factors are perfectly independent, the value of SS_{x1x2} will be zero. The strength of the overall effect of two factors is measured by multiple $\eta^2 = (SS_{x1} + SS_{x2} + SS_{x1x2})/SS_y$. Then, the significance of the overall effect is tested by the following F-test.

$$F = \frac{[(SS_{x1} + SS_{x2} + SS_{x1x2})/(c_1c_2 - 1)]}{[SS_{\text{error}}/(\text{N} - c_1c_2)]} \quad (4)$$

Once we conclude a significant overall effect from the above F-test, we need to examine the significance of the main effect of each of the two factors. Specifically, the significance test of the main effect of X_1 is based on the following F-statistics.

$$F = \frac{[SS_{x1}/(c_1 - 1)]}{[SS_{\text{error}}/(\text{N} - c_1c_2)]} \quad (5)$$

The significance test for the main effect of X_2 can be conducted in the same way. Furthermore, we need to test the significance of the interaction effect of the two factors involved in the analysis. For the null hypothesis of no interaction effects, the appropriate F-test is

$$F = \frac{[SS_{x1x2}/\{(c_1 - 1)(c_2 - 1)\}]}{[SS_{\text{error}}/(\text{N} - c_1c_2)]} \quad (6)$$

Analysis of Covariance (ANCOVA)

ANCOVA is a method of statistical analysis devised by Ronald Fisher in 1932 that combines regression analysis with ANOVA. The procedure involves measuring one or more concomitant variables called as *covariates* in addition to the dependent variable. ANCOVA includes at least one categorical independent variable (factor) and at least one metric-independent variable (covariate) along with one dependent

variable. The most common use of covariates is to remove extraneous variation from the dependent variable because the effects of the factors are of primary interest. For example, in determining how consumers' intentions to buy a brand (dependent variable) vary with different levels of price (factor), age difference (covariate) may have to be taken into consideration. As another example, in determining how multiple groups exposed to different commercials (factor) evaluate the brand advertised (dependent variable), we may want to control for consumers' prior brand knowledge (covariate).

A covariate represents a source of variation that has not been controlled in the experiment and one that is believed to affect the dependent variable (Kirk, 1995). In ANCOVA, the dependent variable can be adjusted to remove the effects of the uncontrolled source of variation represented by the covariates. Its primary advantages are twofold: (i) increase statistical power and (ii) reduce bias (Glass and Hopkins, 1984). That is, ANCOVA is used to produce (i) a reduction in error variance arising from covariates, hence, increased power and (ii) a reduction in bias caused by differences among subjects, where those differences are not attributable to the variation of the independent variable.

Multivariate Analysis of Variance (MANOVA)

MANOVA is an extension of ANOVA to accommodate more than one dependent variable. It is a technique that measures the differences for two or more metric-independent variables based on a set of categorical independent variables. It is different from the N-way ANOVA that accommodates more than one independent variable but with a single dependent variable. Therefore, ANOVA is called a *univariate procedure* because it is used to assess group differences on a single metric-dependent variable. In contrast, MANOVA is known a *multivariate procedure* because it is used to assess group differences between multiple metric-dependent variables simultaneously. That is, in MANOVA, each treatment group is observed on two or more dependent variables (Hair *et al.*, 1995).

In ANOVA, the null hypothesis is that the means of the dependent variable are equal across the category groups defined by independent variables involved. The null hypothesis of MANOVA is that the vectors of means on multiple dependent variables are equal across those category groups. MANOVA can be applied when there are two or more dependent variables that are correlated with one another. If there are uncorrelated multiple dependent variables, we can simply use ANOVA for each dependent variable, respectively.

As an example of MANOVA, let us suppose that four consumer groups, each consisting of 100 randomly selected consumers, are exposed to four different commercials about Tide detergent (Malhotra, 2007). After viewing the commercial assigned to the group, each consumer provides his or her own ratings on preference for Tide, preference for Procter & Gamble (the company marketing Tide), and preference for the commercial itself. Because these three preference measures (three metric-dependent variables) are expected to be highly correlated, MANOVA should be conducted to determine the commercial that is the most effective in producing the highest preferences across the three dependent variables. The ANOVA procedure, which can be applied to each dependent variable separately in this example, cannot account for the strong correlations among the three dependent measures.

Comparison of ANOVA and Multiple Regression

ANOVA and MULTIPLE REGRESSION are two of the most popular statistical analysis tools in marketing research (*see* MARKETING RESEARCH PROCESS). For this reason, they are frequently compared in various ways. On the one hand, the two approaches have much in common, and, accordingly, are categorized into the general linear model. Furthermore, both approaches share the same purpose of explaining the variation in the metric-dependent variable associated with the variation in the independent variables.

On the other hand, comparisons of ANOVA and multiple regression often emphasize differences between the two analysis tools in terms of examining the relationships between the

dependent variable and independent variables (Kirk, 1995). Notably, the two approaches differ in the nature of the independent variables used. Specifically, in the typical regression application, the researcher is interested in explaining a metric-dependent variable using one or more metric-independent variables. By contrast, in the typical ANOVA application, the researcher is interested in the categorical independent variables that can explain the variation in the metric-dependent variable. However, this difference is blurred when both approaches are extended to accommodate the other type of independent variables. Regression analysis is not restricted to the use of metric-independent variables and can accommodate categorical or nonmetric variables in the form of dummy variables (e.g., males vs. females; buyers vrs. nonbuyers). By the same token, ANCOVA, an extended analysis tool of ANOVA, can include metric-independent variables as covariates when they are nuisance variables that need to be controlled for. Even so, ANOVA cannot measure the impact magnitude of each independent variable on the dependent variable as regression analysis does.

From a practical standpoint, ANOVA is appropriate in situations where the independent variable is set at certain specific treatment levels and metric measurements of the dependent variable are obtained at each of the treatment levels considered. Therefore, researchers often rely on ANOVA to analyze data from marketing experiments employing rigorous research designs. For that reason, ANOVA is commonly used in consumer behavior research that often requires a high level of control of the experiment process. For instance, ANOVA would be a proper analysis tool in an experiment in which we select different groups of stores to examine various amounts of in-store promotion for a brand and measure results on the brand sales as the dependent variable (Parasuraman, Grewal, and Krishnan, 2004). In contrast, multiple regression is commonly used in econometric models for marketing problems that use secondary data because it can measure the magnitude of influence of each independent variable on the dependent variable. This type of research does not intend to control the data collection process tightly as with usual consumer behavior research.

Applications of Analysis of Variance (ANOVA) in Marketing Research

ANOVA can be applied to cases where the researcher is interested in examining the varying effects of categorical independent variables (e.g., different MP3 player brands) on the metric-dependent variable (e.g., consumer satisfaction in music listening). Since the analysis technique has been widely applied in marketing research, particularly in consumer behavior, most marketing research books discuss the technique in varying degrees of detail with practical application examples. Three such examples are briefly introduced here.

First, Cooper and Schindler (2006) use the example of CalAudio, an MP3 manufacturing company. In this example, the marketing manager intends to investigate whether the quality of their MP3 players is affecting their customers' repurchase decisions negatively because their closest competitor's product appears to have fewer repair problems leading to higher customer satisfaction ratings. The manager decides to monitor two measures to assess product quality: adherence to product specifications and time before failure. The specification variable is measured on a 0-to-100 scale, with 100 meeting all product specifications. The time-before-failure variable is calculated in weeks. Furthermore, the management asks the industrial engineering department to develop a modified manufacturing procedure that can improve the product quality measures without decreasing the production rate. After testing the new modified manufacturing procedure, the marketing manager takes a sample of 15 MP3 players made with the old manufacturing method and 15 made with the new method. The players are measured on the two dependent variables – their adherence to product specification and their time before failure. This example is the case of MANOVA because there are two dependent variables (i.e., product specifications and time before failure) with one factor (old vs. new manufacturing procedures).

On the other hand, Malhotra (2009) uses a one-way ANOVA example of Levi Strauss & Company. In his example, the stiff competition in the jeans market forces the company consider three alternative image themes of the

jeans (e.g., "Levi's. They go on"; "They're not Levi's jeans until we say they are"; and "A style for every story"). The researcher assigns a group of respondents to each of the three image themes (independent variable with three categories) and asks each group to evaluate its matched theme by rating it on a 10-point scale (metric-dependent variable).

Similarly, Hair *et al*. (2008) introduce another one-way ANOVA example based on the Santa Fe Grill database. In this case, the owners want to know whether customers who come to the restaurant from greater distances differ from customers who live closer in their willingness to recommend the restaurant to their friends. In this example, there is a single metric-dependent variable (i.e., customers' willingness to recommend) with two levels of a single factor (customers from greater distances vs. customers who live closer).

Bibliography

Churchill, G.A. and Iacobucci, D. (2005) *Marketing Research: Methodological Foundations*, Thomson South-Western.

Cooper, D.R. and Schindler, P.S. (2006) *Marketing Research*, McGraw-Hill Irwin.

Feinberg, F.M., Kinnear, T.C., and Taylor, J.R. (2008) *Modern Marketing Research: Concepts, Methods, and Cases*, Atomic Dog.

Glass, G.V. and Hopkins, K.D. (1984) *Statistical Methods in Education and Psychology*, Allyn and Bacon.

Hair, J.F. Jr., Anderson, R.E., Tatham, R.L., and Black, W.C. (1995) *Multivariate Data Analysis*, 4th edn, Pearson Prentice Hall.

Hair, J.F., Wolfinbarger, M., Ortinau, D.J., and Bush, R.P. (2008) *Essentials of Marketing Research*, McGraw-Hill Irwin.

Kirk, R.E. (1995) *Experimental Design: Procedures for the Behavioral Sciences*, 3rd edn, Brooks/Cole Publishing Company.

Malhotra, N.K. (2007) *Marketing Research: An Applied Orientation*, 5th edn, Prentice Hall.

Malhotra, N.K. (2009) *Basic Marketing Research: A Decision-Making Approach*, 3rd edn, Prentice Hall.

McDaniel, C. and Gates, R. (2008) *Marketing Research Essentials*, John Wiley & Sons, Inc.

Parasuraman, A., Grewal, D., and Krishnan, R. (2004) *Marketing Research*, Houghton Mifflin Company.

cluster analysis

Michel Wedel and Wei Shi (Savannah)

Introduction

Cluster analysis is a collection of methods for categorizing consumers into clusters that are homogeneous along a range of variables (demographics, attitudes, lifestyles, perceptions). In marketing, it is most often applied for purposes of market segmentation, and the variables that are used to derive the segments are called *segmentation bases*. Other important applications occur in data mining and in perceptual mapping (where products instead of consumers are clustered). An important distinction is between hierarchical and nonhierarchical methods of cluster analysis. Hierarchical methods are "bottom-up" procedures that start with each subject in his/her own cluster and arrive at groupings by successively aggregating the individual clusters. Nonhierarchical methods could be called *top down* and start with some initial allocation of subjects to two or more clusters and employ various algorithms to improve those initial clusters.

Hierarchical Clustering Methods

The methods in this class all begin with each consumer in an individual cluster. Then, they seek subjects that are similar, and join them to form a cluster. Similarity is defined on the basis of on a (dis)similarity metric between the consumers being clustered. Hierarchical classification methods result in a dendrogram: a tree structure that represents the hierarchical relations among all subjects being clustered. Figure 1 shows an example of such a dendrogram, with consumers labeled A through L.

We can distinguish branches in the tree: the horizontal lines; nodes at which the branches join; and the clusters, groups of subjects joined together in a subtree. The horizontal axis displays the distance between consumers, on the basis of a distance metric defined by the user when running the clustering algorithm. For each node in the graph (where a new subcluster is formed), we can read off the distance at which its elements were linked together into a new cluster, on the horizontal axis. Note that the tree itself does not define the clusters. One needs to "cut" the tree at certain levels to get clusters of consumers. For example, cutting the tree in Figure 1 at a distance of 20 produces two clusters.

Similarity measures. Most hierarchical clustering methods use dissimilarities or distances between consumers when forming the clusters. Those measures assess how similar or dissimilar the consumers to be clustered are. They are derived from the variables measured on the consumers. The choice of the type of similarity measure used is made by the researcher, and will depend on whether the variables are discrete or continuous as well as on the type of similarity that is desired from a substantive point of view. While there are many such measures, the two most important of these are discussed here.

Continuous variables. The *Euclidian distance* is the geometric distance that one usually computes between two points in space. Assume that two variables are denoted by X_{i1} and X_{i2} have been measured on consumer i. Then, the measure of how (dis)similar consumers i and j are is obtained by plotting them as points, for example, (X_{i1}, X_{i2}) and (X_{j1}, X_{j2}) in the plane spanned by these two variables, and computing the distance between them as

$$d_{ij} = \sqrt{(X_{i1} - X_{j1})^2 + (X_{i2} - X_{j2})^2} \qquad (1)$$

The Euclidian distance is thus the square root of the squared differences of the coordinate values of the two points that represent the consumers. For two variables, one can visually plot these two points, but this is a bit more difficult for three variables, and impossible for four or more. However, even in those cases, the Euclidian distance can still be computed: in the case of P variables being measured, the sum in Equation 1 is extended to a sum of P (squared) coordinate differences. A related distance measure is the *Mahalanobis distance*, which corrects for the differences in scales and covariance between the variables. If the covariance matrix is the identity matrix, the Mahalanobis distance reduces to the Euclidean distance.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

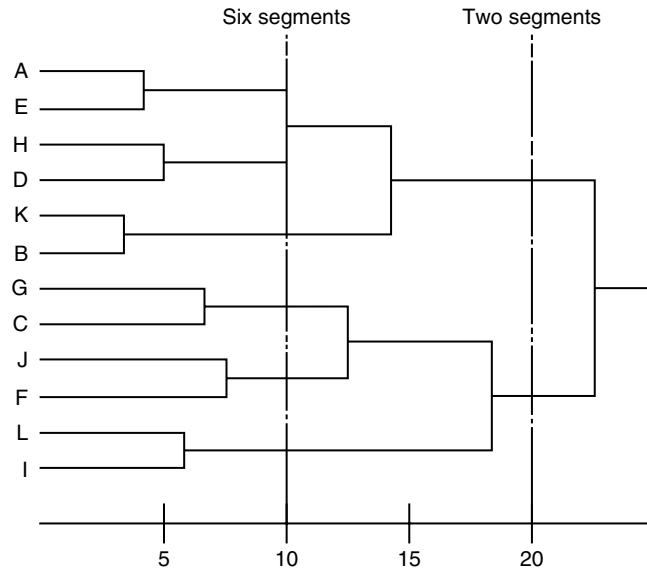


Figure 1 A dendrogram resulting from hierarchical clustering.

A second measure that is often used for continuous variables, such as lifestyle ratings, is the *Manhattan distance*:

$$d_{ij} = |X_{i1} - X_{j1}| + |X_{i1} - X_{j1}| \qquad (2)$$

the sum of the absolute values of the differences of the coordinate values. Again, for P variables, there would be a sum of P terms in Equation 2. Instead of taking the distance of the shortest line between two points in a higher dimensional space, as the Euclidian distance does, the Manhattan distance measures distance by "walking" perpendicular along the coordinate axes, hence its name.

These distance measures need to be computed for all pairs of consumers, leading to $N(N-1)/2$ distances among N consumers. This is so since the distance from subject i to j is the same as that from j to i, while the distance from subject i to itself is 0. Thus for 10 subjects, there would be 45 distances. The distance matrix has 0 as its diagonal elements, is symmetric, and has nonnegative off-diagonal elements. We could use Euclidian distance, Mahalanobis distance, or Manhattan distance for this matrix.

Table 1 Cross classification of two consumers, i and j, on a number $(a + b + c + d)$ binary variables.

i/j	*Yes*	*No*
Yes	*a*	*b*
No	*c*	*d*

Discrete variables. For discrete variables, different (dis)similarity measures need to be computed. Again, while there are many, we discuss two in detail. Assume that there are several binary variables (for example, yes/no readership of 25 magazines) assessed for consumers i and j. Then, a measure of how similar the subjects are on the basis of reading interest could be how often they both do or do not read a magazine. We can cross-classify the two consumers as shown in Table 1.

In Table 1, the number a denotes how often the two share a "yes," d denotes how often they share a "no," b and c show how often one reads a magazine and the other does not. Clearly, the higher that a and d are relative to the total, the more similar the consumers may be. Thus, the *simple matching coefficient* of similarity

between consumers i and j is defined as

$$S_{ij} = \frac{a+b}{a+b+c+d} \tag{3}$$

Jaccard's matching coefficient is defined as

$$S_{ij} = \frac{a}{a+b+c} \tag{4}$$

A measure of dissimilarity is obtained (similar to the Euclidian distance, for example) if one computes $1 - S_{ij}$.

The two are similar, but differ in that Jaccard's coefficient does not count joint absence to contribute to similarity, which may be reasonable in some applications. Both Equations 4 and 5 vary between 0 and 1, denoting that the subjects are completely dissimilar and completely similar, respectively.

Again, these distance measures are computed for all $N(N-1)/2$ pairs of consumers to obtain a distance matrix.

Clustering algorithms. The similarities or distances among subjects constitute the input to hierarchical clustering algorithms. Agglomerative hierarchical methods seek to cluster consumers by agglomerating groups of consumers at successive stages of the algorithm, starting with single-consumer clusters. Several algorithms are available for that purpose. All algorithms start with joining subjects that have the smallest distance (or the largest similarities) into clusters of two consumers. But, they differ in the way they define the distance between a consumer and a cluster, or between two clusters, in later stages of the algorithm. While such definitions are to some extent arbitrary, they do affect the types of clusters of subjects that are obtained.

At each stage of the *single linkage* (also called *nearest neighbor*) cluster algorithm, the distance between two clusters is defined as the *smallest* distance between any two members of the two clusters. The *complete linkage* algorithm (also called *furthest neighbor*) uses the *largest* distance between any two members of the two clusters to decide which clusters are closest. In *average linkage*, the distance between two clusters composed of several consumers is defined as the *average* distance between all pairs of consumers in the two clusters. In *centroid linkage*, the distance between two clusters is defined as the *Euclidian distance between the centers* of the clusters. Cluster center is computed as the average of the variables for all consumers within the cluster. If the cluster center is instead defined as the median, or most representative consumer in the cluster, *median linkage* is obtained. Figure 2 illustrates all the above cluster algorithms. *Ward's method* uses the *variance* of all observations in two clusters as a measure of the distance between these two clusters. In other words, if all consumers in two clusters combined have a large variance, the distance between these two clusters is large; if they have a small variance, the distance between the two clusters is small, and these clusters will be merged in earlier stages of the algorithm.

The aforementioned distance measures can be used to calculate the *agglomerative coefficient*. It

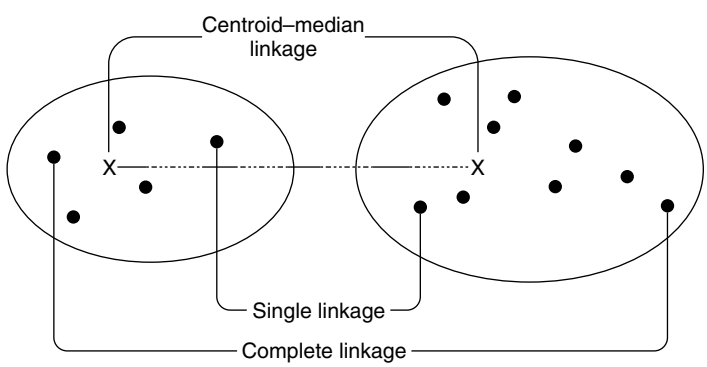


Figure 2 Illustration of the definition of distance between clusters for various hierarchical clustering algorithms.

indicates the extent to which an existing cluster and the next closest cluster or variable will be combined. We illustrate the relationship between agglomerative coefficient and number of clusters in Figure 3.

Nonhierarchical Clustering Algorithms

The nonhierarchical clustering methods do not derive a tree structure from the data. Rather, they partition the data into a predetermined number of segments, minimizing some criterion of interest. The most popular nonhierarchical algorithm, *K-means* clustering, seeks to minimize the (Euclidian) distances between all consumers within each cluster. The K-means procedure starts by assuming a certain number of clusters, say two. It then assigns all subjects to one of the two clusters, and since this is just a way to start the algorithm, a random assignment will do. Alternatively, the algorithm could be initialized by assigning the two consumers that are farthest apart as the initial cluster centers, and then allocating the remaining consumers to the closest center.

Subsequently, K-means computes the centroids of the two clusters by computing the average of the variables across the subjects in each cluster. Thus, if there are, for example, two brand-personality variables on which consumers are to be clustered, one obtains two mean brand-personality scores for each cluster $(\overline{X}_{c1}, \overline{X}_{c2})$, with $c = 1$ and $c = 2$ denoting the clusters. Then, the algorithm moves consumers from one cluster to another, and assigns a consumer to the cluster with the closest centroid. The distance measure of a consumer to a cluster centroid that is used is the Euclidian distance

$$d_{ic} = \sqrt{(X_{i1} - \overline{X}_{c1})^2 + (X_{i2} - \overline{X}_{c2})^2} \qquad (5)$$

Then, the algorithm computes the centroids of the clusters again. The algorithm tries all possible transfers of all consumers from one

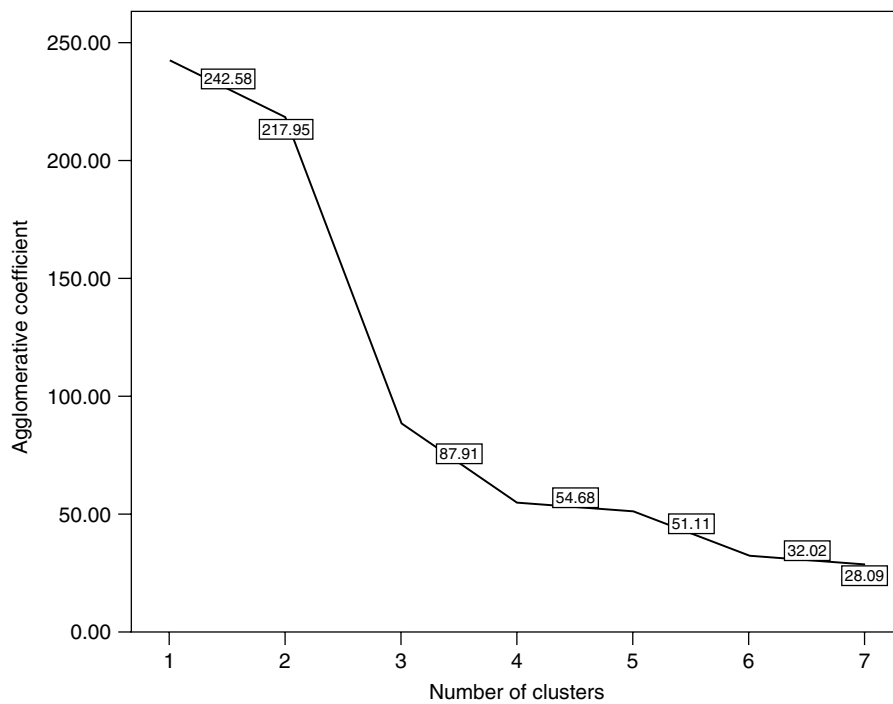


Figure 3 An illustration of the relationship between agglomerative coefficient and number of clusters. Source: SPSS Tutorial/Cluster Analysis, on judges.sav data.

cluster to the other, and keeps on cycling through all consumers until no more improvement is achieved. Again, this can be extended to P variables by extending Equation 5 to a sum of P terms.

A variant of the K-means algorithm is the *partitioning around mediods* (PAMs). The PAM algorithm works similarly to K-means, but instead of computing the centroids (means of the variables) it computes the mediods, which are the variable values of the most "representative consumer" for each cluster. The mediod is the consumer with the smallest average distance to all other consumers in the cluster in question.

The K-means and PAM algorithms provide a "hard assignment" of all consumers to clusters. In other words, a specific consumer is either member of a cluster 1 or is not. *Fuzzy clustering algorithms* extend this by allowing consumers to have only part membership in each of the clusters, and assign membership values that are between 0 and 1 for each consumer, where the sum of these values is 1. The fuzzy c-means algorithm, for example, works similarly to the K-means, but instead of reassigning consumers to clusters at each stage of the algorithm, it recomputes their fuzzy membership value in each cluster.

Number of Clusters

In hierarchical methods, one way to identify the optimal number of clusters is to inspect the dendrogram, and hope to find particularly long branches. Such a long branch may indicate that the cluster below it is relatively homogeneous, and therefore one would like to cut the tree at those long branches, yielding two segments, for example, as in Figure 1.

In nonhierarchical methods, the number of clusters is assumed before the analysis (it was assumed to be two above). However, the optimal number of clusters to represent a particular dataset is unknown. One usually tries several values, for example, 2, 3, . . ., 10, and looks at a criterion value to see which solution is the best, which is often conveniently done by plotting the criterion value against the number of clusters. Such a criterion usually is a measure of the within-cluster variability, such as the average within-cluster distance between subjects (respectively, the percentage of variance explained in the data by the cluster solution). While these measures will always decrease (respectively, increase) if the number of clusters increases, one hopes to find an "elbow" in the curve, yielding a solution after which the decrease (respectively, increase) levels off, or is relatively flat. For hierarchical algorithms, such a plot may be obtained as well, by plotting the number of clusters against the amalgamation coefficient, which is the numerical value of the distance measure used by the specific clustering algorithm, at which clusters merge.

One powerful criterion for nonhierarchical methods is the average *silhouette distance*, which is a measure of the extent to which subjects are misclassified (positive silhouette length values indicate correct classification, negative misclassification). The solution with the larger silhouette distance is the better one, and the silhouette distance should be larger than 0.25. For clustering algorithms, for which a likelihood measure can be computed, standard information statistics such as AIC (Akaike information criterion), BIC (Bayesian information criterion), or CAIC (consistent Akaike information criterion) can be used to select the number of clusters.

Relative Performance of Clustering Methods

Among the hierarchical methods discussed here, average linkage and Ward's method are reported to perform best. Single linkage tends to be less stable, and is subject to the emergence of long and ill-separated clusters, a phenomenon called "*chaining*." One problem in applying the hierarchical clustering methods is the availability of a large number of dissimilarity measures, distances, and clustering algorithms, and the choice among them is often subjective. This holds to a lesser extent for the nonhierarchical methods, which optimize a specific criterion of cluster homogeneity. The nonhierarchical methods are reported to perform better than the hierarchical methods: they are less sensitive to outliers in the data and to irrelevant variables, that is, variables that are present in the data but do not contribute to the cluster solution. However, the nonhierarchical methods are sensitive to the starting solution that is

chosen, and different clustering solutions may be obtained depending on how subjects are allocated to the clusters when starting the algorithm (this is called the *problem of "local optima"*). This holds in particular for large datasets. A hierarchical algorithm may be used to obtain a starting solution for a nonhierarchical algorithm, which tends to improve performance.

Cluster Validation

In applications of cluster analysis, cluster validation is important. The final cluster solution may strongly depend on the subjective choice of variables used as a basis for clustering, distance measures, clustering algorithms, and procedures to select the optimal number of clusters. This makes cluster validation essential. There are several ways to validate a cluster solution. The first is *significance testing*. We expect to see the clusters differ in new variables that were not used in the clustering. Often, (multivariate) analysis of variance is used for this purpose, with cluster membership as the independent variables. A note of caution is that testing the variables themselves that were used as input for clustering is inappropriate, and the significance tests between clusters become invalid.

A second powerful way to validate a cluster solution is through *replication*. The dataset is (randomly) split in two or more parts, and the cluster analysis is replicated on the two parts. Consumers in part 2 are assigned to the clusters in part 1, and one checks this corresponds to the clustering of part 2 itself (for example, using cross tabulation). This yields a measure of internal consistency of the cluster solution.

Finally, it is recommended to check *consistency* of the cluster solutions by repeating the clustering with different distance measures and clustering algorithms, and checking whether the solutions obtain correspond (again, cross tabulation can be used).

Cluster Analysis Software

Software for cluster analysis is widely available. The major statistical packages, including SPSS, S-Plus, BMDP, SAS, STATA, and Minitab have a collection of clustering routines. In addition, R has build-in functions as well as contributed packages for clustering.

For related topics, see LATENT CLASS AND FINITE MIXTURE MODELS.

There is an extensive literature on cluster analysis. Most books on multivariate analysis, or marketing research contain chapters on cluster analysis, for example, Hair *et al.* (1995). The booklet by Aldenderfer and Blashfield (1984) provides a quick yet accurate introduction. Everitt (1993) provides a concise yet condensed description of clustering routines. The book by Rousseuw and Kaufman (1990) is advanced and detailed, and describes the powerful algorithms that are implemented in the S-Plus and R clustering packages. Wedel and Kamakura (2000) describe clustering methodology and its applications in market segmentation.

Bibliography

Aldenderfer, M.S. and Blashfield, R.K. (1984) *Cluster Analysis*, Sage Quantitative Applications in the Social Science, Number 07-044, Sage Publications, London.

Everitt, B.S. (1993) *Cluster Analysis*, Edward Arnold, London.

Hair, J.F., Anderson, R.E., Tatham, R.L., and Black, W.C. (1995) *Multivariate Data Analysis*, Prentice Hall, Englewood Cliffs.

Rousseeuw, P.J. and Kaufman, L. (1990) *Finding Groups in Data*, John Wiley & Sons, Inc., New York.

SPSS, Inc. (2001) The SPSS Twostep Cluster Component. SPSS white papers/technical report TSCPWP-0101, SPSS, Chicago.

Wedel, M. and Kamakura, W.A. (2000) *Market Segmentation: Conceptual and Methodological Foundations*, Kluwer, Dordrecht.

conjoint analysis

Vithala R. Rao

Introduction

Several interdependent decisions are involved in the formulation of a marketing strategy for a brand (of a product or service). These include decisions on not only the product's characteristics but also its positioning, communication, distribution, and pricing to chosen sets of targeted customers. The decisions will need to be made in the wake of uncertain competitive reactions and a changing environment. For a business to be successful, the decision process must include a clear understanding of how customers will choose among (and react to) various competing alternatives. It is well accepted in marketing that choice alternatives can be described as profiles on multiple attributes and that individuals consider various attributes while making a choice. While choosing, consumers typically make trade-offs among the attributes of a product or service. Conjoint analysis (CA) is a set of techniques ideally suited to studying customers' choice processes and determining trade-offs.

CA is probably the most significant development in marketing research methodology over the last 40 years or so. The method has been applied in several thousand applied marketing research projects since its introduction to the marketing researchers in 1971 (Green and Rao, 1971). The method has been applied successfully for tackling several marketing decisions such as optimal design of new products, target market selection, pricing a new product, and competitive reactions. A significant advantage of the method has been the ability to answer various "what if" questions using market simulators; these simulators are based on the results from a conjoint study for hypothetical and real choice alternatives.[1]

Five different features of CA have contributed to its versatility for tackling marketing managerial problems: (i) it is a measurement technique for quantifying buyer trade-offs and values; (ii) it is an analytical technique for predicting buyers' likely reactions to new products/services; (iii) it is a segmentation technique for identifying groups of buyers who share similar trade-offs/values; (iv) it is a simulation technique for assessing new product service ideas in a competitive environment; and (v) it is an optimization technique for seeking product/service profiles that maximize share/return (Green, Krieger, and Wind, 2004).

Against this brief background, this article will be organized as follows. In the next (second) section, principal types of CAs that are in vogue in marketing research are described. In the third section, the process for conducting a conjoint study is briefly described; this section also includes a discussion of various data collection formats and designs for developing stimuli (or profiles) for a conjoint research problem. The basics of conjoint models and estimation are described in the fourth section and a simplified illustration of one approach is provided. In the fifth section, an overview of the variety of applications of this method is presented. A series of recent developments and future directions are enumerated in the sixth section with limited elaboration.

Principal Types of Conjoint Analysis

Over the past several years, various researchers have contributed to the general methodology of CA. The reader is referred to Green and Srinivasan (1978, 1990) for excellent reviews of the field of CA. Essentially, there are four types of conjoint methods; the traditional method (CA) that uses stated preference ratings; the choice-based conjoint analysis (CBCA) that uses stated choices; the adaptive conjoint analysis (ACA) developed in part to handle the issue of a large number of attributes; and the self-explicated CA, which is a bottom-up method. The first three of these can be called *decompositional methods* because the stated preference or stated choice data are decomposed to obtain part-worth functions. The fourth one is called the *compositional method* because it composes a preference score from ratings of scores on attribute levels and relative importances of attributes. We briefly describe each of these.

The traditional CA collects preferences (judgments) for profiles of hypothetical products each described on the entire set of attributes

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

selected for the conjoint study. These profiles are called *full profiles*. However, when one concatenates levels of all attributes, the complete set of full profiles (or full factorial design), will in general, be very large. A respondent will be unduly burdened when asked to provide preference judgments on all profiles. Typically, a smaller set of full profiles (selected according to an experimental design) is used in a conjoint study. An individual's overall stated preferences are decomposed into separate and compatible utility values corresponding to each attribute typically using regression-based methods. These separate functions are called *attribute-specific part-worth functions*. In most cases, the preference functions can be estimated at the individual level. This estimated preference function can be deemed as an indirect utility function.

While the traditional decompositional conjoint approach involved full profiles of product concepts described on multiple attributes, several new data collection formats have emerged over the years. A significant development is the use of data on stated choices elicited under hypothetical scenarios that mimic the marketplace and estimating part-worth functions from such data using primarily multinomial logit methods (*see* LOGIT MODEL); these methods are labeled choice conjoint methods (CBCA or CBC) and have become popular in the early 1990s and are probably the most widely used methods currently. They are based on the behavioral theory of random utility maximization McFadden, 1974; the origin of this approach is the law of comparative judgment developed by Thurstone (1927). This approach decomposes an individual's random utility for an object into two parts: deterministic utility and random part. Depending on the distributional assumptions for the error part, a number of alternative models are developed to describe the probability of choice of an object. The most popular one is the multinomial logit model (MNL) (*see* LOGIT MODEL) that uses the extreme value distribution for the error term. These methods belong to the family of discrete choice analysis methods. An excellent volume that elaborates on these stated choice methods is by Louviere, Hensher, and Swait (2000); see also Ben-Akiva and Lerman (1991).

Researchers have also developed an adaptive conjoint method which is called *ACA* (Johnson, 1987). The method involves first a self-explicated task (i.e., eliciting data on attribute importances and attribute level desirabilities using ranking and subsequent rating) followed by preference ratings for a set of partial profiles descriptions, two at a time using a graded, paired comparison scale. The partial profile descriptions are tailored to each respondent based on the data collected in the self-explicated task. Both the tasks are administered by computer. This method is a type of hybrid[2] model approach.

In contrast, the compositional approach based on the multiattribute attitude models (Wilkie and Pessemier, 1973) estimates preferences from judged values of the components (importances and desirabilities) that contribute to preference. In the compositional approach, individuals are asked to evaluate the desirability of each level of all the attributes as well as the relative importances assigned to the attributes. Then, the preference for any product concept is estimated as a weighted sum of the desirabilities for the specific levels of attributes describing that concept; the weights are the relative importances. This approach is called the *"self-explicated" method* (see Green and Srinivasan, 1978 for more details). Studies have shown that the self-explicated method is surprisingly quite robust (Srinivasan and Park, 1997).

Process of Conducting a Conjoint Study

In order to provide a context for designing a conjoint study, consider the problem of determining the steady-state demand for a new product. Assume that a sample of n consumers drawn randomly from a total of N customers in the target market for this product was interviewed in a conjoint study. Let q_i denote the quantity of product bought by ith customer in the sample; $i = 1, 2, \ldots, n$ (generally measured in the survey) and let p_i denote the probability that the ith customer will purchase the new product in a steady state (conditional on his/her consideration set of alternative items including the new product). Then, the demand forecast

for the new product is given by the model:

$$D = \left(\frac{N}{n}\right) \sum_{i=1}^{n} q_i p_i$$

The problem then is to estimate the probability of purchase p_i for the new product for the members of the sample. There are at least two solutions for this problem.

One solution is to employ the traditional stated preference-based CA. This method involves obtaining an estimate of the utility a customer derives for a new product described in terms of its attributes relative to other items considered and then translating the utility into probabilities of purchase. Several methods exist for this transformation; see Green and Krieger (1988). The second solution is to utilize the CBCA method. In this approach, the probabilities can be computed directly from the MNL model. Against this backdrop, the process of designing conjoint studies is reviewed.

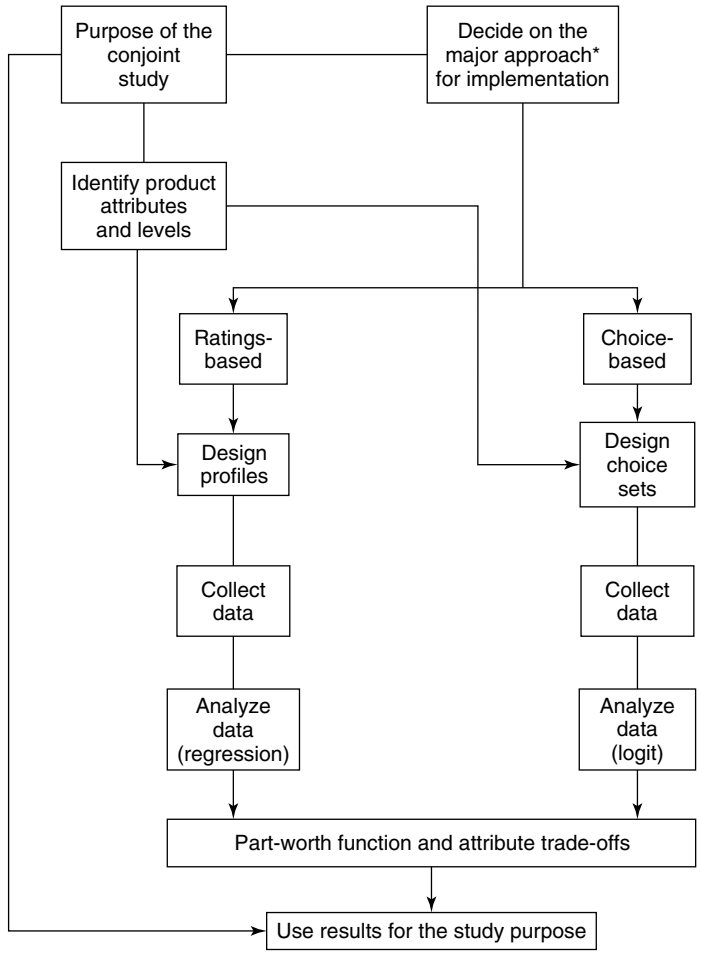
A typical CA project for collecting and analyzing stated preference or stated choice data[3] as such consists of four main steps: (i) development of stimuli based on a number of salient attributes (hypothetical profiles or choice sets); (ii) presentation of stimuli to an appropriate sample of respondents; (iii) estimation of part-worth functions for the attributes as well as any heterogeneity among the respondents; and the use of the estimates in tackling any managerial problems (e.g., forecasting, pricing, or product design). The steps are schematically shown in Figure 1.

One major step is the design of stimuli (either profiles or choice sets). To illustrate profiles and choice sets, consider a simple conjoint problem with three attributes, A, B, and C, each described at four levels, as a1, a2, a3, and a4. An example profile is (a2, b3, c1) and an example choice set is {(a1, b2, c3); (a2, b1, c4); (a4, b3, c2); (No choice)} with some times "no choice" not included. Stated preference for a profile is measured as a rating or a rank relative to other profiles, while stated choice is the choice made by the respondent among the items in a choice set.

The aspect of designing stimuli (profiles or choice sets) has received much focus since the beginning of CA; it draws much from the theory of experimental design, where procedures for constructing subsets of combinations of all attribute levels are developed. CA for ratings-based studies makes extensive use of orthogonal arrays (Addelman, 1962; Green, 1974). The process for designing choice sets for collecting stated choice data is a lot more complicated; after developing a number of profiles (usually a subset of all possible profiles), subsets of profiles (4 or 5) are used as choice sets. The "no choice" option is generally included in each choice set. Street and Burgess (2007) present the theory and methods for the construction of optimal stated choice experiments; see also Street, Burgess, and Louviere (2005) and Street and Burgess (2004). Researchers can use the OPTEX procedures in the SAS system (2002–2003) for designing profiles or choice sets; see also Kuhfeld (2005).

In the ratings-based conjoint approach, the respondent is given a number of profiles of product concepts, each described on the attributes under study, and is asked to state his/her preference for each profile on a rating scale (e.g., 10 points or 100 points). These preference data are analyzed using multiple regression methods (typically, a dummy variable ordinary least squares (OLS) regression) to estimate a utility function for each respondent (or for a subgroup of respondents). We illustrate this approach in the next section. Typically, additive utility functions are used although utility functions with interaction terms are possible depending on the experimental designs used for constructing profiles.

The attributes in a conjoint study are of two types: categorical or continuous (or interval-scaled) with only a few selected values (*see* MODELS FOR CATEGORICAL DEPENDENT VARIABLES). A categorical attribute (such as low, medium, or high) can be converted into a number of dummy variables (one less than the number of levels). A continuous attribute (such as price of a product) can be used directly or can also be converted into dummy variables; if used directly, only a linear term or with both linear and quadratic terms to account for any nonlinear effects can be included in the utility function. With suitable redefinitions of variables, the utility function for the ratings methods can be written as $y = X\beta + \varepsilon$, where



*Several alternatives exist here; two are highlighted.
Source: Rao, V. R. "Developments in conjoint analysis" in B. Wierenga (ed.) *Handbook of marketing decision models*, New York: Springer science + business media, LLC, 2008.

Figure 1 Major steps in a conjoint study (reproduced from B. Wierenga (Ed) Handbook of Marketing Decision Models, © Springer Science + Business Media, LLC, 2008).

ε is the random error of the model assumed to be normally distributed with zero mean and variance of σ^2, y is the rating on given profile, and X is the corresponding set of p dummy (or other) variables. The model is estimated using regression methods (usually ordinary least squares method; *see* MULTIPLE REGRESSION). The β is a $px1$ vector of regression coefficients associated with the dummy variables or continuous variables included in the model. The part-worth values for each attribute can be derived from these regression coefficients.

In the choice conjoint methods, the respondent is given a number of choice sets, each choice set consisting of a small number (typically 4 or 5) of profiles, and is asked to indicate which profile will be chosen. An MNL is used for estimating the deterministic component of the random utility using maximum likelihood methods. A variety of extensions and alternatives exist for

analyzing stated choice data. The MNL model (*see* LOGIT MODEL) for the choice conjoint data is the probability of choosing profile j in choice set $C = \exp(-v_j)/\sum \exp(-v_k)$, where the summation is taken over all the profiles in the choice set C and v_j is the deterministic component of the utility for the profile j. The deterministic utility function v is specified analogous to the linear combination to the function for y in the ratings methods. The estimated coefficients will be used in computing the part-worth values for the attributes in the study.

Current approaches for implementing a CA project differ in terms of several features; some main features are stimulus representation, formats of data collection, nature of data collection, and estimation methods. Table 1 lays out some alternatives for these features. There is no clear agreement as to which data collection format is the best and all the formats shown in Table 1 are in vogue.

One notable development is the use of hierarchical Bayesian estimation methods (*see* UNOBSERVED HETEROGENEITY) which enable an analyst to incorporate prior knowledge in the part-worth values as monotonic or other types of order constraints in the estimation process (Allenby, Arora, and Ginter, 1995); see also Lenk *et al.* (1996). Further, part-worth functions are estimated at the aggregate (or subgroup) level or at an individual level. Researchers have also used finite mixture methods (DeSarbo *et al.*, 1992) to "uncover" segments of respondents based on the preference or choice data collected in conjoint studies; see also Andrews, Ansari, and Currim (2002). The variety of recently developed techniques for estimation of part-worth functions is very impressive and a discussion of these is beyond the scope of this article. For a recent discussion of conjoint methods, see Hauser and Rao (2004) and Rao (2008).

Basics of Conjoint Models

Conjoint methods are intended to "uncover" the underlying preference function of a product in terms of its attributes[4]. A general product profile defined on r attributes can be written as

Table 1 Alternatives for selected features of conjoint analysis.

Representation of Stimuli	*Formats of Data Collection*	*Nature of Data Collection*	*Estimation Methods*
Verbal descriptions	Full profile evaluations	One-shot Adaptive	Regression-based methods
Pictorial descriptions	Partial profile evaluations	Multiple times[a]	Random utility models
Videotapes and supporting materials	Stated preferences		Direct computation based on self-explicated importances
	Stated choices		Hierarchical Bayes estimation[a]
Virtual prototypes			
	Self-explicated methods		
Combinations of physical models, photographs and verbal descriptions	Configurators[a]		Methods based on new optimization methods[a] Analytic center estimation, support-vector machines, genetic algorithms

[a] These are newer methods.
Source: Adapted from Hauser and Rao (2004) and Rao (2008).

$(x_{j1}, x_{j2}, \ldots, x_{jr})$, where x_{jt} is the level for the jth profile on the tth attribute in a product profile. While there exist several ways for specifying the preference functions in CA, researchers usually start with an additive conjoint model; but, the theory extends to models with interactions as well. The preference score[5] for the jth product profile, y_j for one respondent additive conjoint model is

$$y_j = U_1\left(x_{j1}\right) + U_2\left(x_{j2}\right) + \cdots + U_r\left(x_{jr}\right)$$

where $U_t(\bullet)$ is the component utility function specific to the tth attribute (also called *part-utility function* or *part-worth function*). No constant term is specified, but it could be included in any one of the U-functions or assumed to be zero (without any loss of generality). The specification of the U-function for any attribute will depend upon its type (categorical and quantitative). In practice, a conjoint study may contain both types of attributes.

Brand names or verbal descriptions such as high, medium, or low are examples of a categorical attribute; here the levels of the attribute are described by words. A quantitative attribute is one measured by either an interval scale or a ratio scale; numbers describe the "levels" of such an attribute; examples are the weight of a laptop and speed of the processor.

The levels of a categorical attribute can be recoded into a set of dummy variables and a part-worth function is specified as a piecewise linear function in the dummy variables. In this case, the component utility function for a categorical attribute (tth for example) will be

$$U_t\left(x_{jt}\right) = D_{t1}U_{t1} + D_{t2}U_{t2} + \cdots + D_{tr_t}U_{tr_t}$$

where r_t is the number of discrete levels for the tth attribute (resulting from the construction of the profiles or created ex post); D_{tk} is a dummy variable taking the value 1 if the value x_{it} is equivalent to the kth discrete level of x_t and 0 otherwise; and U_{tk} is the component of the part-worth function for the kth discrete level of x_t. In practice, only $(r_t - 1)$ – one less the number of discrete levels of the attribute – dummy variables are necessary for estimation.

A quantitative attribute can be used in a manner similar to a categorical attribute by coding its values into categories or used directly in the specification of the part-worth function for the attribute. In the latter case, the function can be specified as linear (vector model) or nonlinear; one example of a nonlinear function is the ideal point model. Mathematically, the component utility function can be specified as

$$U_t\left(x_{jt}\right) = \begin{cases} w_t x_{jt} & \text{for the vector model; and} \\ w_t\left(x_{jt} - x_{0t}\right)^2 & \text{for the ideal point model;} \end{cases}$$

where w_t is a weight (positive or negative) and and x_{0t} is the ideal point on the tth attribute.

A linear function is appropriate for an attribute deemed to be desirable (e.g., speed of a laptop computer) or undesirable (e.g., weight of a laptop computer); such a function is called a vector model for which the utility increases (or decreases) linearly with the numerical value of the attribute.

As mentioned above, with suitable redefinitions of variables, the preference function can be written as $y = X\beta + \varepsilon$, where ε is the random error of the model assumed to be normally distributed with zero mean and variance of σ^2; y is the rating on a given profile; and X is the corresponding set of p dummy (or other) variables. The β is a $px1$ vector of part-worths among the levels of attributes.

At this point, it will be useful to indicate the software available for designing and implementing conjoint studies. These are

- Sawtooth Software (ACA, CBC, etc.; probably the most complete solution)
- SPSS (useful for preference-based approach)
- SAS (OPTEX for design and several other programs for analysis)
- LIMDEP (useful for analyzing data of various types; Greene, 2003)
- Bayesm package in R (developed by Rossi, Allenby, and McCulloch, 2005)
- MATLAB (one needs to develop a specific program code).

Table 2 Variable smart phone attributes.

Attribute	*Levels*			
	1	*2*	*3*	*4*
Style	Candy bar	Slide phone	Flip phone	Touch screen
Brand	Blackberry	Nokia	LG	Samsung
Talk time	3 h	5 h	7 h	9 h
Weight	100 g	115 g	130 g	145 g
Camera quality	2 MP	3 MP	6 MP	8 MP

An Illustration of Ratings-based Conjoint Analysis

As an illustration of the ratings-based conjoint method, assume that a wireless provider firm is interested in determining trade-offs among various features of a smart phone (a technologically advanced product with a number of features). In order to simplify the data collection, assume that the firm is interested in the trade-offs among five attributes, namely, style of the phone, brand name, talk time, weight, and camera quality (having predetermined a number of standard features). Price attribute was not included because it was part of a contract with the wireless provider and was about the same for all brands. Each of these five attributes is varied at four levels. Table 2 shows the features that are predecided and the levels of the five features varied in the study.

The total number of possible hypothetical profiles are 1028 ($= 4 \times 4 \times 4 \times 4 \times 4$), which are combinations of the levels of the five attributes. Given that it is almost impossible to have a respondent judge all these profiles, the study designer has selected 32 of these profiles using a fractional factorial design. In the study, respondents were shown the complete list of standard features and were asked to provide preferences on a 0–100 point scale for the 32 profiles. These profile descriptions were provided using a computerized questionnaire. The results from an analysis of one respondent's evaluations are shown. These data are analyzed using dummy variable regression after converting each attribute into three dummy variables as shown in Table 3 to obtain part-worth functions for the five attributes. The resulting regression and part-worth values are also shown in Table 3. The range of each part-worth function is a simple measure[6] of the importance of that attribute. Figure 2 shows the plots of part-worth functions for the five attributes.

Although not shown, the fit of the part-worth model to the individual's preference ratings is quite good with an adjusted R-square of 0.88. On the basis of this analysis, one can conclude that this respondent has a strong preference for a flip style Blackberry smart phone that is lightest in weight with a talk time of 9 hours and a camera quality of 6 MP. From the graphs

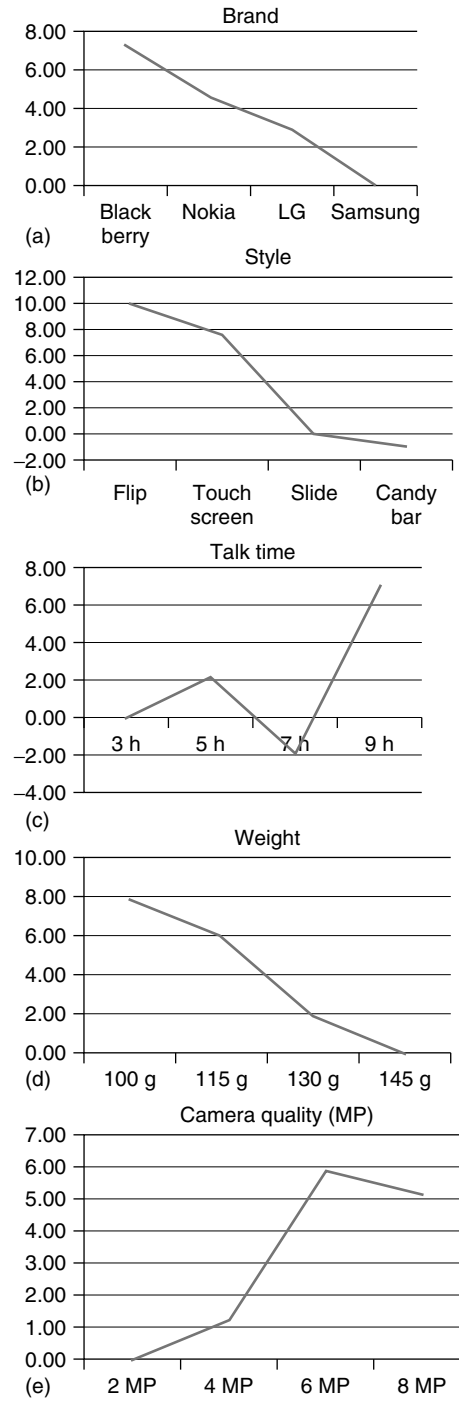


Figure 2 Part-worth functions for the smart phones illustration weight.

Table 3 Estimated part-worth values for the illustrative data for smart phone study.

Attribute	*Level*	*Recoded Dummy Variables*			*Part-worth Value*	*Relative Importance(%)*
		D1	*D2*	*D3*		
Brand	Blackberry	1	0	0	7.25	17.86
	Nokia	0	1	0	4.63	
	LG	0	0	1	2.87	
	Samsung	0	0	0	0.00	
Style	Candy bar	1	0	0	−0.97	26.42
	Flip	0	1	0	9.95	
	Touch screen	0	0	1	7.58	
	Slide	0	0	0	0.00	
Talk time	3 h	0	0	0	6.99	21.84
	5 h	0	0	1	−1.88	
	7 h	0	1	0	2.12	
	9 h	1	0	0	0.00	
Weight	100 g	1	0	0	7.88	19.4
	115 g	0	1	0	6.01	
	130 g	0	0	1	1.89	
	145 g	0	0	0	0.00	
Camera quality	2 MP	0	0	0	0.00	14.48
	4 MP	1	0	0	1.22	
	6 MP	0	1	0	5.88	
	8 MP	0	0	1	5.15	

of the part-worths, one can see that the decline in utility from these levels of the attributes to other levels is not uniform. Further, there is nonlinearity in the part-worth function for the attribute of the camera quality. Looking at the relative importances, this individual places most importance for the style attribute followed by talk time, weight, brand, and camera quality in that order. The part-worth functions can be sued to predict the individual's preference rating for a profile not covered in the 32 profiles. Further, one can estimate preferences for items in any choice set; these estimates can be used to predict the individual's first choice and other choices. Also, if the study is conducted for a sample of respondents, the vector of estimated relative importances can be used to form clusters of individuals whose importances are quite similar; these clusters are akin to market segments. Focusing on one brand (e.g., LG), one can make predictions of first and other choices for the sample for various scenarios (e.g., anticipated changes in the product designs of competing brands); such a process is the simulation aspect of CA, which is highly useful for managers.

Applications

Since its introduction, conjoint methods have been applied in a large number of applied marketing research projects. There is no recent estimate of the number of applied studies but its use is increasing tremendously. The conjoint methodology has been applied in several areas; these include consumer nondurable products (bar soaps, carpet cleaners, lawn chemicals, etc.), industrial goods (copying machines, portable computer terminals, personal computer design, etc.), other products (car batteries, ethical drugs, pesticides, etc.), financial services (branch bank services, auto insurance policies, credit card features, etc.), transportation (domestic airlines, electric car design, etc.), and other services (hotel design, car rental agencies, telephone pricing, etc.). The method has been applied

successfully for tackling several marketing decisions such as the optimal design of new products, target market selection, pricing a new product, and studying competitive reactions. Some high profile applications of these techniques include the development of Courtyard Hotels by Marriott (Wind *et al.*, 1989) and the design of the E-Z Pass Electronic Toll Collection System in New Jersey and neighboring States in the United States (Green, Krieger, and Vavra, 1997). A significant advantage of the conjoint method has been the ability to answer various "what if" questions using market simulators; these simulators are based on the results of an analysis of conjoint data collected on hypothetical and real choice alternatives.

Some Recent Developments

We have mentioned the development of hierarchical Bayesian methods and experimental design described earlier in the article. In addition, there have been developments on dealing with a positive part-function for price (Rao and Sattler, 2003), use of incentive-aligned methods for data collection (Ding, Grewal, and Liechty, 2005; Ding, 2007), a range of methods for handling large number of attributes (reviewed in Rao, Kartono, and Su, 2008), polyhedral methods aimed at reducing respondent burden (Toubia *et al.*, 2003; Toubia, Hauser, and Simester, 2004), and modeling choices for bundles (Bradlow and Rao, 2000; Chung and Rao, 2003) and upgrading methods (Park, Ding, and Rao, 2008) based on the BDM method (Becker, DeGroot, and Marschak, 1964), experimental designs based on new criteria such as utility balance (Huber and Zwerina, 1996; Hauser and Toubia, 2005), continuous CA (Wittink and Keil, 2003; Su and Rao, 2006), adaptive self-explicated analysis (Netzer and Srinivasan, 2007), and measuring reservation prices for single products and bundles (Jedidi and Zhang, 2002; Jedidi, Jagpal, and Manchanda, 2003). These are but only a few examples of continuous developments in CA research. The article written from the 2007 Choice Symposium, Netzer *et al.* (2008) identifies several new directions in this methodology; see also Hauser and Rao (2004), Bradlow (2005), and Rao (2008) for ideas for future research in this area. In conclusion, one might say that CA is alive and well!

Endnotes

[1] It will be useful to review some terms used in conjoint analysis. Attributes are (mainly) physical characteristics that describe a product; levels are the number of different values an attribute takes; profile is a combination of attributes, each attribute at a particular level, presented to a respondent for an evaluation (or stated preference); choice set is a pre-specified number of profiles presented to a respondent to make a pseudo-choice (stated choice).

[2] Hybrid models involve a combination of several tasks aimed to increase the "efficiency" of data collection in conjoint studies usually for large number of attributes. See Green (1984) for a review of these methods; see also Green and Krieger (1996). We will not delve much into these methods due to space limitations.

[3] For the sake of ease in exposition, we will restrict to these two types of data and will not delve into methods that involve variations such as the hybrid methods.

[4] For an introduction to conjoint analysis, see Orme (2006).

[5] For exposition purposes, I am considering a ratings-based conjoint analysis where respondents provide preference ratings for a number of product profiles. The same can apply to the v-function in the choice-based conjoint analysis.

[6] There are several other measures such as partial R-squared but are beyond the scope of this chapter.

Bibliography

Addelman, S. (1962) Orthogonal main-effect plans for asymmetrical factorial experiments. *Technometrics*, **4**, 21–46.

Allenby, G.M., Arora, N., and Ginter, J.L. (1995) Incorporating prior knowledge into the analysis of conjoint studies. *Journal of Marketing Research*, **37**, 152–162.

Andrews, R.L., Ansari, A., and Currim, I. (2002) Hierarchical Bayes versus finite mixture conjoint analysis models: a comparison of fit, prediction, and partworth recovery. *Journal of Marketing Research*, **39**, 87–98.

Becker, G.M., DeGroot, M.H., and Marschak, J. (1964) Measuring utility by a single-response sequential method. *Behavioral Science*, **9**, 226–232.

Ben-Akiva, M. and Lerman, S.R. (1991) *Discrete Choice Analysis*, MIT Press, Cambridge.

Bradlow, E.T. (2005) Current issues and a 'Wish List' for conjoint analysis. *Applied Stochastic Models in Business and Industry*, **21**, 319–323.

Bradlow, E.T. and Rao, V.R. (2000) A Hierarchical Bayes model for assortment choice. *Journal of Marketing Research*, **37**, 259–268.

Chung, J. and Rao, V.R. (2003) A general choice model for bundles with multiple category products: application to market segmentation and optimal pricing for bundles. *Journal of Marketing Research*, **40**, 115–130.

DeSarbo, W.S., Wedel, M., Vriens, M., and Ramaswamy, V. (1992) Latent class metric conjoint analysis. *Marketing Letters*, **3** (3), 273–288.

Ding, M. (2007) An incentive-aligned mechanism for conjoint analysis. *Journal of Marketing Research*, **44**, 214–223.

Ding, M., Grewal, R., and Liechty, J. (2005) Incentive-aligned conjoint analysis. *Journal of Marketing Research*, **42**, 67–82.

Farquhar, P.H. and Rao, V.R. (1976) A balance model for evaluating subsets of multiattributed items. *Management Science*, **22**, 528–539.

Green, P.E. (1974) On the design of choice experiments involving multifactor alternatives, *Journal of Consumer Research*, **1**, 61–68.

Green, P.E. (1984) A hybrid models for conjoint analysis: an expository view. *Journal of Marketing Research*, **21**, 33–41.

Green, P.E. and Krieger, A.M. (1988) Choice rules and sensitivity in conjoint analysis. *Journal of Academy of Marketing Science*, **6** (Spring), 114–127.

Green, P.E. and Krieger, A.M. (1996) Individualized hybrid models for conjoint analysis. *Management Science*, **42**, 850–867.

Green, P.E., Krieger, A., and Vavra, T.G. (1997) Evaluating new products. *Marketing Research*, **9**, (4) 12–21.

Green, P.E., Krieger, A., and Wind, Y. (2004) Buyer choice simulators, optimizers, and dynamic models, in *Marketing Research and Modeling: Progress and Prospects: A Tribute to Paul E. Green* (eds Y. Wind and P.E. Green), Kluwer Academic Publishers, Norwell.

Green, P.E. and Rao, V.R. (1971) Conjoint measurement for quantifying judgmental data. *Journal of Marketing Research*, **8**, 355–363.

Green, P.E. and Srinivasan, V. (1978) Conjoint analysis in consumer research: issues and outlook. *Journal of Consumer Research*, **5**, 103–123.

Green, P.E. and Srinivasan, V. (1990) Conjoint analysis in marketing: new developments with implications for research and practice. *Journal of Marketing*, **54**, 3–19.

Greene, W.H. (2003) *Econometric Analysis*, 5th edn. Pearson Company.

Hauser, J.R. and Rao, V.R. (2004) Conjoint analysis, related modeling, and applications, in *Marketing Research and Modeling: Progress and Prospects: A Tribute to Paul E. Green* (Y. Wind and P.E. Green), Kluwer Academic Publishers, Norwell.

Hauser, J.R. and Toubia, O. (2005) The impact of utility balance and endogeneity in conjoint analysis. *Marketing Science*, **24** (3), 498–507.

Huber, J. and Zwerina, K. (1996) On the importance of utility balance in efficient designs. *Journal of Marketing Research*, **33**, 307–317.

Jedidi, K., Jagpal, S., and Manchanda, P. (2003) Measuring heterogeneous reservation prices for product bundles. *Marketing Science*, **22** (1), 107–130.

Jedidi, K. and Zhang, Z.J. (2002) Augmenting conjoint analysis to estimate consumer reservation price. *Management Science*, **48** (10), 1350–1368.

Johnson, R.M. (1987) Adaptive conjoint analysis, in *Sawtooth Software Conference on Perceptual Mapping, Conjoint Analysis, and Computer Interviewing*. Sawtooth Software, Inc., Ketchum, pp. 253–265.

Johnson, R.M. (1991) Comment on "Adaptive conjoint analysis: some caveats and suggestions". *Journal of Marketing Research*, **28**, 223–225.

Kuhfeld, W.F. (2005) *Marketing Research Methods in SAS: Experimental Design, Choice, Conjoint, and Graphical Techniques*, SAS 9.1 Edition, TS-722, SAS Institute Inc., Cary.

Lenk, P.J., DeSarbo, W.S., Green, P.E., and Young, M.R. (1996) Hierarchical Bayes conjoint analysis: recovery of partworth heterogeneity from reduced experimental designs. *Marketing Science*, **15** (2), 173–191.

Liechty, J.C., Fong, D.K.H., and DeSarbo, W.S. (2005) Dynamic models incorporating individual heterogeneity: utility evolution in conjoint analysis. *Marketing Science*, **24** (2), 285–293.

Louviere, J.J., Hensher, D.A., and Swait, J. (2000) *Stated Choice Methods*, Cambridge University Press, Cambridge, New York.

McFadden, D. (1974) Conditional logit analysis of qualitative choice behavior, in *Frontiers in Econometrics* (ed. P. Zarembka), Academic Press, New York, pp. 105–142.

Netzer, O. and Srinivasan, V. (2007) Adaptive-self Explication of Multi-attribute Preferences, Working Paper, Graduate School of Business, Stanford University, Stanford.

Netzer, O., Toubia, O., Bradlow, E.T. *et al.* (2008) Beyond conjoint analysis: advances in preference measurement. *Marketing Letters*, **19**, 337–354.

Orme, B. (2006) *Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research*, Research Publishers LLC, Madison.

Park, Y-H., Ding, M., and Rao, V.R. (2008) Eliciting preference for complex products: a web-based upgrading method. *Journal of Marketing Research*, **45**, 562–574.

Rao, V.R. (2008) Developments in conjoint analysis, in *Handbook of Marketing Decision Models* (ed. B. Wierenga), Springer, pp. 23–55.

Rao, V.R., Kartono, B., and Su, M. (2008) Methods for handling massive number of attributes in conjoint analysis. *Review of Marketing Research*, **5**, 104–129.

Rao, V.R. and Sattler, H. (2003) Measurement of price effects with conjoint analysis: separating informational and allocative effects of price, in *Conjoint Measurement: Methods and Applications*, 3rd edn (eds A. Gustafsson, A. Herrmann, and F. Huber), Springer, Berlin.

Rossi, P.E., Allenby, G.M., and McCulloch, R. (2005) *Bayesian Statistics and Marketing*, John Wiley & Sons, Ltd, West Sussex, England.

SAS Institute (2002–2003) *Statistical Analysis System*, SAS Institute, Cary, NC.

Srinivasan, V. and Park, C.S. (1997) Surprising robustness of self-explicated approach to customer preference structure measurement. *Journal of Marketing Research*, **34** (2), 286–291.

Street, D.J. and Burgess, L. (2004) Optimal and near-optimal pairs for the estimation of effects in 2-level choice experiments. *Journal of Statistical Planning and Inference*, **118**, 185–199.

Street, D.J. and Burgess, L. (2007) *The Construction of Optimal Stated Choice Experiments: Theory and Methods*, John Wiley & Sons, Inc., New York.

Street, D.J., Burgess, L., and Louviere, J.J. (2005) Quick and easy choice sets: constructing optimal and nearly optimal stated choice experiments. *International Journal of Research in Marketing*, **22**, 459–470.

Su, M. and Rao, V.R. (2006) A Continuous Conjoint Analysis for Preannounced New Products with Evolutional Attributes, Working paper, Johnson School, Cornell University.

Thurstone, L. (1927) A law of comparative judgment. *Psychological Review*, **34**, 273–286.

Toubia, O., Hauser, J.R., and Simester, D.I. (2004) Polyhedral methods for adaptive conjoint analysis. *Journal of Marketing Research*, **42**, 116–131.

Toubia, O., Simester, D.I., Hauser, J.R., and Dahan, E. (2003) Fast polyhedral adaptive conjoint estimation. *Marketing Science*, **22** (3), 273–303.

Wilkie, W.L. and Pessemier, E.A. (1973) Issues in marketing's use of multi-attribute attitude models. *Journal of Marketing Research*, **10** (4), 428–441.

Wind, Y., Green, P.E., Shifflet, D., and Scarbrough, M. (1989) Courtyard by Marriott: designing a hotel with consumer-based marketing. *Interfaces*, **19**, 25–47.

Wittink, D R. and Keil, S K. (2003), Continuous conjoint analysis, in (eds. A. Gustaffsson, A. Herrmann and F. Huber), Conjoint Measurement: Methods and Applications, Springer, Berlin, pp. 541–564.

exploratory factor analysis

Michel Wedel and Wei Shi (Savannah)

Factor Analysis and Principal Components Analysis

Factor analysis is not a single method, but in fact a group of methods with unprecedented popularity in marketing research. It is applied as a "data-reduction" or as a "latent-structure" detection method. In the first application we are in the domain of data mining, in which factor analysis continues to be a popular and powerful tool. The second type of application often occurs in market structuring based on subjective judgments, and in perceptual mapping. The main purposes of factor analysis are (i) to reduce the number of variables in a dataset, (ii) to explore relationships among those variables, and (iii) to detect latent dimensions (factors) underlying the relationships between observed variables. Two sets of goals, (i) and (ii) and (ii) and (iii), correspond to two different variants of factor analysis: principal components analysis (PCA) and factor analysis in the strict sense, respectively.

The Basic Idea

Suppose we want to measure satisfaction with an administrative computer system, among a sample of $i = 1, \ldots, N$ consumers. We design a satisfaction questionnaire in which we ask our consumers how satisfied they are with hardware reliability (X_{i1}), with software reliability (X_{i2}), with the ease of maintaining the hardware (X_{i3}), and with the ease of maintaining the software (X_{i4}). Most likely, the responses to the four questions are correlated with each other. From relatively high correlations among the variables, we can conclude that there is a certain degree of redundancy between them. Therefore, one may wish to combine the original variables into a single new one that captures the essential information. One way to do this is to sum the responses to all questions, for each consumer. Let us call this new variable a *factor* (or alternatively a *component*; please note that the word "factor" is used in different ways in statistics: it can also be used to indicate a categorical variable with a finite number of levels, such as education). The score for consumer i on the new factor is obtained from the following equation:

$$PC_i = X_{i1} + X_{i2} + X_{i3} + X_{i4} \qquad (1)$$

We have achieved parsimony, since instead of the four original variables, we now have one new variable, or factor, and we have the "score" of each consumer on that new factor, or principal component (PC is an abbreviation of principal component, another term for factor).

The question is, however, whether summing is the best way to combine the observed X-variables. For example, the redundancy between the reliability items (X_{i1} and X_{i2}) may be large and the redundancy between the maintenance questions (X_{i3} and X_{i4}) may be large, but the redundancy between these two sets may be limited. In other words, consumers' scores on reliability items are correlated, and those on maintenance questions are correlated. Alternatively, there might be redundancy (correlation) between the hardware (X_{i1} and X_{i3}) and between the software (X_{i2} and X_{i4}) variables respectively, but not between those two groups of variables. Thus, a better way to combine those variables could be to use weights to obtain the scores on the new factor:

$$PC_i = W_1X_{i1} + W_2X_{i2} + W_3X_{i3} + W_4X_{i4} \qquad (2)$$

Here, instead of summing the variables, a weighted combination of the variables is computed, with four weights, one for each variable: W_1, W_2, W_3, W_4. The question, of course, is how can we get those weights? We can set these weights manually, according to a priori notions of how the variables are to be combined. The simpler Equation 1 is obtained by setting all weights to 1. Alternatively, we could construct a new factor as in Equation 2, by summing all hardware questions (set their weights to 1, weights for software to 0). Or we could construct a "software factor" by summing all software questions (set their weights to 1, weights for hardware to 0).

We could also choose to set the weights based on the correlations between the variables, where

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

those variables that correlate highly are weighted together. This has the advantage that we let the data tell us the variables that are to be combined into the factor, rather than making the decision ourselves a priori. For example, if there is much redundancy among the questions X_{i1} and X_{i2} of reliability of software and hardware, as evidenced by a high correlation among them, we would want to give X_{i1} and X_{i2} high weights (i.e., W_1, W_2 are large). However, since now the maintenance variables X_{i3} and X_{i4} receive low weights, we get little information on them. But if these two variables are highly correlated as well, we may wish to form a second factor, independent of the first, that has large weights for the two maintenance variables (i.e., W_3, W_4 are large). Thus, we now have two equations, one for each factor:

$$PC_{i1} = W_{11}X_{i1} + W_{21}X_{i2} + W_{31}X_{i3} + W_{41}X_{i4}$$
$$PC_{i2} = W_{12}X_{i1} + W_{22}X_{i2} + W_{32}X_{i3} + W_{42}X_{i4} \quad (3)$$

As mentioned above, if we want the first factor to be only a combination of X_{i1} and X_{i2} and the second factor to capture the information in X_{i3} and X_{i4}, we could set $W_{31} = W_{41} = W_{12} = W_{22} = 0$. Although we have two factors, we have obtained parsimony, since, instead of the original four variables, we now have two new factors that we can work with. In general, having the data tell us what the weights should be is a better idea.

If we have a large number of variables, we may want to compute more than two factors. The problem then becomes somewhat more complicated, since we need to choose many weights based on the correlations between the variables. In the above case of four variables and two factors, we have $2 \times 4 = 8$ weights to choose; in the general case of $\mathcal{J}$ variables and P factors we have $\mathcal{J} \times P$ weights. The problem is complicated even more if we do not know beforehand how many factors we want to extract (for example, in the example above it might not be clear whether we want to combine the variables based on reliability–maintenance, software–hardware distinctions, or both). PCA can do this for us: it estimates the weights based on the correlation of the variables, and helps us to choose the correct number of factors to be derived from the data.

Principal Component Analysis

As mentioned above, the purpose of PCA is data reduction. We have observed $\mathcal{J}$ variables, but would like to work with P components instead ("components" is used rather than "factors" in this section, which is just a matter of conforming to the standard terminology). P is smaller than $\mathcal{J}$. We call the principal components approach *formative*, since it "forms" the components from the observed variables as in Equation (3). Later, we will distinguish this from the "reflective" factor analysis procedure. PCA is not a statistical model, but a data-analysis or data-mining tool to reduce the data and explore meaningful patterns. This will be explained in a bit more detail below.

Estimation of the model is not discussed here. It suffices to mention that it is based on a mathematical technique called *eigenvalue decomposition*. This method has the interesting property that, based on the correlation matrix of the variables as input, it provides us with those components that explain the maximum percentage of variance in all original variables (similar to the way a linear regression explains the maximum amount of variance attainable in a single y-variable). Note, however, that PCA does not distinguish between X- and Y-variables. Equation 3 looks similar to a regression equation, but there are important differences. First, a key difference is that not only the weights W_{ip} but also the "dependent variables" PC_{ip} are unknown. Second, there are multiple "dependent variables" PC_{ip} that are modeled simultaneously. Third, the PCA formulation does not contain error terms (see Equation 3, for example), owing to which it is merely a convenient mathematical representation of the data that does not allow for statistical testing.

An interesting property of the eigenvalue decomposition method is that the components extracted are independent, which makes their interpretation easier than that of the original variables. PCA uses the correlation matrix of the variables as input. The output produced by computer programs that use the eigenvalue decomposition to estimate principal components is discussed below. An example of the output is given in Table 1. Variables are shown in the first column, the columns labeled *Comp#* are the three extracted components, and *Comm* is the

Table 1 Loadings matrix estimated with PCA.

Variable/Factor	*Comp 1*	*Comp 2*	*Comp 3*	*Comm*
Software reliability	0.24	−0.38	0.53	0.48
Hardware reliability	0.27	−0.26	0.48	0.37
Performance	0.30	−0.21	—	0.13
Upgradability	0.27	—	−0.42	0.25
System manageability	0.29	−0.22	—	0.13
Service manageability	0.29	—	−0.34	0.19
Ease of configuration	0.28	—	−0.26	0.14
Ease of maintainability	0.29	—	−0.22	0.13
Servicability	0.28	0.34	—	0.19
Mechanical design	0.27	0.45	—	0.27
Electrical design	0.27	0.44	—	0.26
Software compatibility	0.29	—	—	0.08
Hardware compatibility	0.28	0.28	—	0.16
Eigenvalue	1.00	0.88	0.91	—
Cumulative % explained	0.56	0.65	0.71	—

column with communalities. We use this table to illustrate the main concepts of PCA.

- *Loadings*. Rather than referring to W_{ip} as weights, in PCA one refers to them as *loadings*. They represent the contribution of a variable to a component, and can be interpreted as the correlation between the (unobserved) component and the (observed) variable. Like correlations, loadings can be between minus one and one. For example, the first variable in Table 1, software reliability, has a loading of 0.24 for the first, and −0.38 for the second component. Thus, software reliability correlates positively with the first component, while it has a negative correlation with the second. Thus, consumers that have higher scores for software reliability tend to score higher on component 1, but lower on component 2. The square of the loading is the percentage of the variance in the variable accounted for by the factor: it would be 14.4 and 5.8%, respectively for components 1 and 2, for software reliability.
- *Eigenvalue*. A PCA provides one "eigenvalue" for each component extracted. The term *eigenvalue* derives from the technique used to derive the components (eigenvalue decomposition). That value is proportional to the percentage of variability in all variables explained by that component. In Table 1, the first component has an eigenvalue of 1.00, and explains 56% of the variability in all variables (that percentage is obtained by dividing the eigenvalue by the sum of the eigenvalue of all components).
- *Communality*. The communality represents the percentage of variance in a specific variable, accounted for by all the components and is thus a number ranging from zero to one. It is shown as the rightmost column in Table 1. For example, the PCA explains 48% of the variance in the software reliability variable. It is the sum of the squared loadings in that row of the table; here, it is $0.38^2 + 0.24^2 + 0.53^2 = 0.48$. The communality helps us to inspect to what extent the analysis captures the variability in each of the original variables. Some statistical packages compute *uniqueness* instead of communality. Uniqueness is equal to 1 – communality, so that in the example above it is $1 - 0.48 = 0.52$. Uniqueness is the variance in a particular variable, unexplained by the components. It is that part of the variable which is unique to it, and hence its name.

How to select the number of components. In most applications of PCA, the number of components is unknown. How many components to extract from the data is an arbitrary decision to a certain

extent. In principle, one could extract as many components as there are variables, but in most applications this is not a very useful thing to do because it does not yield any reduction in the data. Here we provide some guidelines that are commonly used, and that, in practice, seem to yield good results.

First, we can retain only components with *eigenvalues greater than* 1. In essence this is like saying that, unless a component extracts at least as much as the equivalent of one original variable, we drop it. The "eigenvalue larger than 1" criterion is probably the one most widely used. In our example above, using this criterion, we have retained one component, since it has an eigenvalue larger than 1. This criterion is a conservative criterion in most applications, since it provides an upper bound to the number of components to retain, and, in most applications, will indicate too large a number of components.

A graphical method to select the number of components is the *scree test*. In this test, one plots the eigenvalues (or equivalently the percentages of variance explained by the components) in a simple line or bar plot against the component number. Figure 1 provides an example (generated with the statistical package S-PLUS). Then, we need to find the place where the smooth decrease of eigenvalues appears to level off when moving from left to right in the plot. To the right of this point, presumably, one finds only "scree" ("scree" is the geological term referring to the debris that collects on the lower part of a rocky slope). Or equivalently, one would look for a jump in this plot. According to this criterion, we would probably retain two or three components in Figure 1. In Figure 1, the horizontal line indicates the "eigenvalue equals to 1," and it is clear that this criterion would indicate a much larger number of components (eight).

Component scores. Once we have estimated the weights or loadings, we can compute each consumer's score on the components, using Equation 3. For example, for a particular consumer, Joe, we substitute in Equation 3 his values $X_{Joe,1}$, $X_{Joe,2}$, $X_{Joe,3}$, and $X_{Joe,4}$ for the

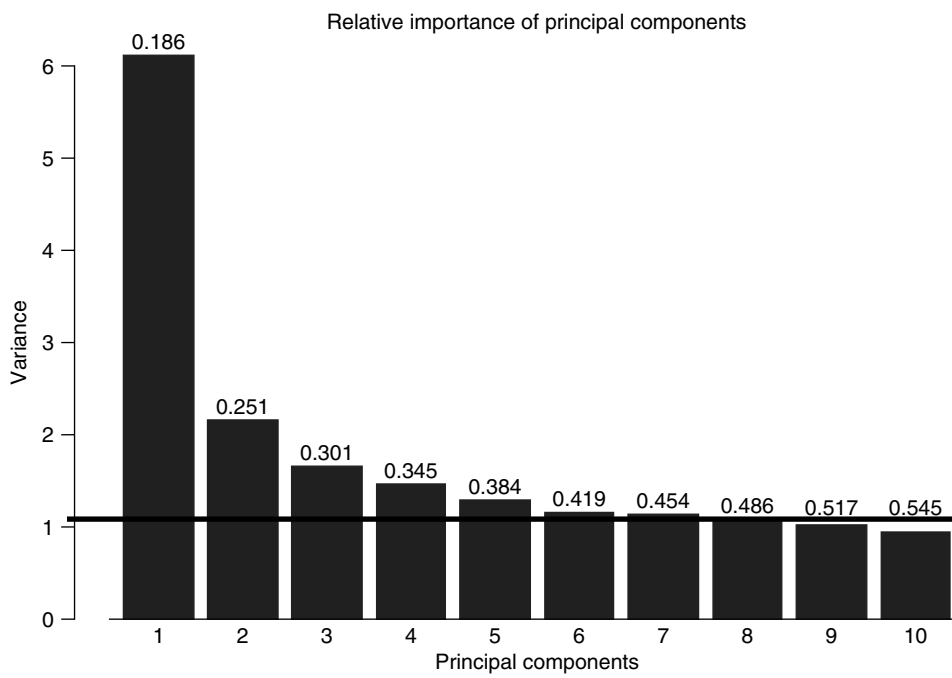


Figure 1 Illustration of a scree plot that plots the relative importance (eigenvalue, variance, and percentage variance explained) of the components.

measured variables, and then substitute the estimated weights $W_{11}, W_{21}, W_{31}, W_{41}$. This gives us a score $PC_{Joe,1}$ for Joe on the first component. For each consumer i, we get a score for each of the P ($P = 2$ in Equation 3) components: PC_{ip}. We can save those scores as (two) new variables in our dataset and use them in subsequent analyses instead of the original variables. A benefit is that these scores are independent. For example, these scores are often used as x-variables in a regression analysis.

Biplot. The biplot is a joint plot of the component loadings and the scores, for two components (usually component 1 and 2, but any other combination of components is possible). Figure 2 provides an example of a biplot, generated with S-PLUS. In that plot, each variable is represented as a line emanating from the origin, and the endpoint of that line is a point that corresponds to the loadings of the variable on components 1 and 2. Each consumer is represented as a point, corresponding to that consumer's scores on components 1 and 2. This allows us to assess the relationships between the variables (variables that are close together in the plot are highly correlated). Also, it enables inspection of the components' scores and their relation to the variables. Two consumers that are close together have similar scores on the components and similar values for the measured variables. Further, one could project the location of a consumer orthogonally on one of the lines representing the original variables. The length of that projection (either positive or negative) tells us whether the consumer scores high on that variable. Thus, the biplot provides a very parsimonious two-dimensional graphical representation of relationships between the rows and columns in the dataset.

Factor Analysis

Factor analysis is related to PCA. But it differs from factor analysis in that it is not a data-reduction tool, but a statistical model, often called a "*measurement model*." In a factor analysis model, the left- and right-hand sides of Equations 3 are reversed:

$$\begin{aligned} X_{i1} &= W_{11}F_{i1} + W_{12}F_{i2} + e_{i1} \\ X_{i2} &= W_{21}F_{i1} + W_{22}F_{i2} + e_{i2} \\ X_{i3} &= W_{31}F_{i1} + W_{32}F_{i2} + e_{i3} \end{aligned} \tag{4}$$

$$X_{i4} = W_{41}F_{i1} + W_{42}F_{i2} + e_{i4} \tag{5}$$

Thus factor analysis specifies a relation between the observed variables ($X_{i1}, \ldots, X_{i4}$), and the unobserved factors (F_{i1}, F_{i2}) and error terms ($e_{i1}, \ldots, e_{i4}$). Note that in Equation 4, the factors as well as the weights on the right-hand side of the equality sign are all unobserved. The model is a statistical model, since it also includes measurement error. It is called a *reflective model*, since it assumes that the unobserved factors are *reflected* in the observed X-variables. This model is often used in measurement problems in marketing, including the measurement of perceptions, attitudes, and so on, as well as in multidimensional scaling, perceptual mapping, and product positioning. The basic idea is that consumers hold only a limited set of perceptions (the factors) that are covered by asking questions on a wide range of observed variables.

Factor analysis and PCA are often confused with one another because the estimation methods for the two models may be similar: the eigenvalue decomposition can be used for both. However, whereas the purpose of PCA is data reduction, that of factor analysis is the identification of latent dimensions. Thus factor analysis requires more theoretical/substantive input: the researcher should know that there are factors, unobserved dimensions, and underlying consumers' responses to a series of measurement scales (often five-, seven-, or nine-point likert scales). However, the researcher does not need to know how many of those factors are there, or what their nature is. If the researcher has precise knowledge on what the number of factors and how they relate to the observed variables, *confirmatory factor analysis* should be applied.

Estimation and number of factors. The eigenvalue decomposition is just one (the simplest) way of estimating factor models; better methods are principal axis factoring (PAF) and maximum likelihood (ML). Both methods are iterative,

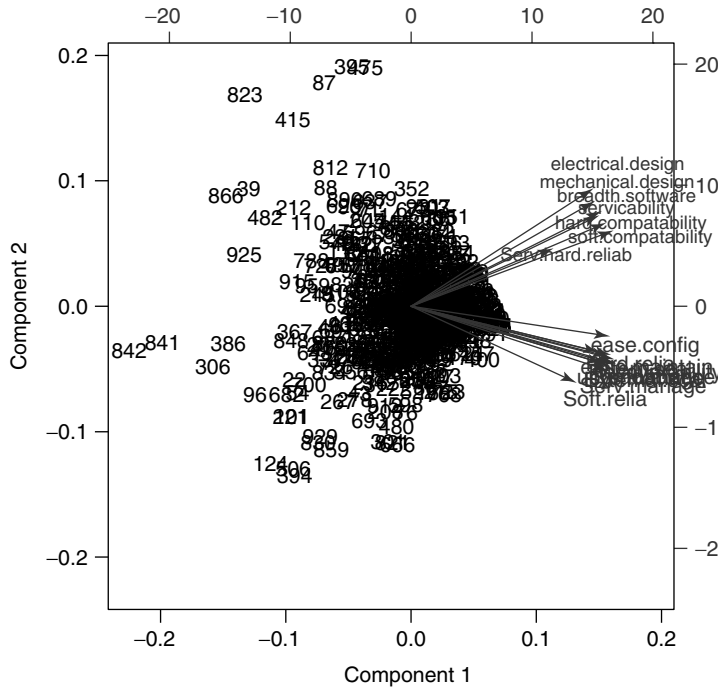


Figure 2 Illustration of a biplot, with the numbers corresponding to consumers and the lines corresponding to measured variables.

but their description is beyond the scope of the present article. ML estimation presupposes that the X-variables all have a normal distribution. This has the advantage that it allows for statistical significance tests for the number of factors and loadings. But in most cases, the number of factors is selected with a scree plot very similar to the one described above for the case of PCA.

Rotation and interpretation of results. We could plot the factor loadings as shown in a biplot (Figure 2). An unrotated factor loading result may be hard to interpret. A common pattern is that nearly all variables have high loadings on the first factor. We solve this problem through rotation. Note that in the biplot, we can rotate the axes in any direction without changing the relative locations of the variables (or the consumers) to each other; however, the actual coordinates of the points, that is, the factor loadings would, of course, change. This rotational indeterminacy is caused by the fact that a factor solution is not unique (that is, there are many different sets of loadings and scores that would give us exactly the same fit to the observe data). That nonuniqueness may help us to obtain a solution that is more easily interpretable. There are various rotation methods that have been proposed. *Varimax Rotation* is the most often used method. The goal of all of these is to obtain a clear pattern of loadings, that is, factors that are somehow clearly marked by high loadings for some variables and low loadings for others. From the rotated factor matrix, the factor loadings can be interpreted or named by the variables that have higher loadings. Such patterns are sometimes referred to as *simple structure*.

Software

Software for PCA and factor analysis is widely available. The major statistical packages, including SPSS, S-PLUS, BMDP, SAS, STATA, and Minitab have a collection of routines and R has build-in functions as well as contributed packages.

For related topics, see the articles CONFIRMATORY FACTOR ANALYSIS; STRUCTURAL EQUATION MODELING.

There is an extensive literature on factor analysis. Most books on multivariate analysis or marketing research contain articles on cluster analysis; see, for example, Hair *et al*. (1995) for an accessible introduction, and Dillon and Goldstein (1984) for a more extensive, yet very clear, exposition.

Bibliography

Hair, J.F., Anderson, R.E., Tatham, R.L., and Black, W.C. (1995) *Multivariate Data Analysis*, Prentice Hall, Englewood Cliffs.

Dillon, W.R. and Goldstein, M. (1984) *Multivariate Analysis*, John Wiley & Sons, Inc., New York.

hypothesis testing related to differences – parametric tests

Marcus J. Schmidt

Decisions Under Uncertainty

Every day we make lots of decisions. The majority of our decisions are of minor importance and the outcomes have limited consequences in our lives. On a given occasion, we decide which retail store we want to shop at, which road to drive on, which clothes to wear, and so on. However, now and then our decisions are important, implying that it is vital for us to make the best of alternating choices. Finding a spouse, choosing a job, purchasing a house, and selecting a school for our child are examples of essential decisions. Many important decisions are related to our work: selecting a supplier, hiring or firing personnel, inventing new technology, changing the production process, and launching a new product. In business settings, it is important that we get things right because making an inappropriate decision may have fatal consequences for the company in terms of economic losses and missed opportunities. Sometimes, wrong management choices even lead to bankruptcy. The financial crises that arose in 2008 may serve as an example.

Since decisions have consequences, management often chooses to use relevant quantitative data as input to the decision-making process. Prior to inventing a new *production process* (*see* PROCESS INNOVATION), management might want to collect empirical support on whether the new process significantly lowers production costs, whether it improves the job satisfaction of employees, and so on. Likewise, before *launching a new product* (*see* LAUNCH STRATEGIES), management might wish to collect empirical support on whether customers are expected to purchase the new product instead of or in addition to the old product or if they may consider switching from the competitor's brand to the company's new brand.

In an ideal world, it is preferable to collect data on all relevant subjects and objects, that is, conduct interviews with all the company's actual and potential customers, record every production run, and so on. However, owing to time and cost considerations, collecting data from the whole universe (N) is rarely possible. Instead, companies, in most cases, base the data collection on a *survey* (*see* SURVEY RESEARCH) where a *sampling technique* (*see* SAMPLING TECHNIQUES) is employed for collecting a *reliable and valid* (*see* RESEARCH RELIABILITY AND VALIDITY) representative subset of N called n.

While limiting data collection to n instead of N has obvious advantages, it also has a disadvantage: sampling error. It may be difficult to prove that n is a valid representation of N.

Assume that management has a well-specified goal concerning customer satisfaction: More than 80% of its 200 000 customers must be either "satisfied" or "totally satisfied" with the company's service. Next, a study carried out by a market research agency among 500 customers shows that 84.3% of customers report to be (totally) satisfied. If we had been interviewing all 200 000 customers, we would have been sure that we had reached our success criteria: our goal would be >160 000 satisfied customers ($0.8 \times 200\,000$) while the empirical data showed 168 600 ($0.843 \times 200\,000$). However, because we have been sampling only 0.25% (500/200 000) of the population, we need to consider possible errors caused by the sampling process (in the following we assume that all errors are chance based).

While 0.843 is greater than 0.80, our empirical sample is only one of an endless number of possible samples. New samples might show satisfaction percentages of, say, 0.741 and even 0.809. What we need is a tool for assessing whether 0.843 based on a sample of 500 cases is significantly greater than 0.80 based on 200 000. Assume that we had conducted 100 random samples each of size $n = 500$. How many would show a satisfaction ratio higher than 0.8? Moreover, how should we react upon the findings? If satisfaction is significantly higher than 0.8, we might drop a planned, expensive promotional campaign. In case it is significantly less than 0.80, we would launch the campaign, ask consultants and sellers to visit less satisfied customers, and so on.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Hypothesis Testing: Testing a Hypothesis about the Mean (Large Sample)

In the first decades of the twentieth century, theorists within the field of scientific research established a formalized procedure. It became common to develop a hypothesis or rather two mutually exclusive hypotheses. Two competing hypotheses are launched, out of which only one represents the true state of nature. The hypothesis under test is normally called the *null hypothesis* and is denoted H_0 while the alternative hypothesis is named H_A. Usually, what we wish to reject is H_0.

An example

H_0: customer satisfaction is less than or equal to 0.8

H_A: customer satisfaction is greater than 0.8

Basically, there are two ways to assess if 0.843 is significantly greater than 0.8 (assuming that the data are normally distributed):

1. We look at the probability value assigned to the empirical measurement. If the probability of empirically observing 0.843 – provided that the true but unobservable value (called μ) is ≤ 0.80 – is less than 0.05 (5 out of 100), then we reject H_0.
2. Another option would be to inspect the length of the so-called *confidence interval*: If the confidence interval associated with 0.843 does not include 0.8, say, that it runs from 0.811 to 0.875, then we reject H_0.

Today it has become common standard for statistical textbooks and for scientific publications to use hypothesis tests based on probability values and confidence intervals. Almost all statistical software (SAS, SPSS, Minitab, etc.) provide a number of quantitative tests.

It should be noted, though, that terms like *hypothesis test*, *probability value*, and *confidence intervals* are of rather recent origin.

The use of probability values was first introduced by the eminent statistician R. A. Fisher (1890–1962). He personally developed most of the significance testing methods now in use (Salsburg, 2001, p. 98). The idea of testing a hypothesis (H_0) against an alternative hypothesis (H_A) was suggested originally by the Polish mathematician Jerzy Neyman (1894–1981) in collaboration with Karl Pearson (1857–1936). Research scientists at that time regarded a hypothesis as either false or true.

In 1934, Karl R. Popper (1902–1994) published *The Logic of Scientific Discovery* in which he argued that scientific statements should be *refutable* or *falsifiable*. Statements or arguments that are vague, imprecise, or qualitative cannot be controlled by way of a statistical test. Such statements are to be characterized as unscientific.

According to Popper there is a need to draw a line in the sand between science and nonscience (metaphysics). He named it *the problem of demarcation*. Popper also argued that even though all empirical evidence appears to support a hypothesis, we can never be 100% sure that it is correct or true. At any point of time, new evidence may refute H_0 and instead support H_1.

Most probably inspired by Popper's falsification concept, research scientists began to develop a more focused approach: a hypothesis (H_0) is either rejected or, alternatively, the effort to reject the hypothesis by way of an alternative hypothesis (H_A) has so far been unsuccessful. Consequently, the null hypothesis is accepted until new empirical evidence leads to its falsification.

One use of tests of significance has been very common. If investigators claim that method A is superior to method B for some purposes, the claim is unlikely to be accepted by their peers in scientific journals unless they can present results of a comparative study that show A superior to B in a test of significance at the 5 or 1% level. This practice has been beneficial in preventing false claims from spreading and in saving the time and experience needed to show later that the claims were false, though the practice presumably has discouraged some work that would have been fruitful.

An empirical test of a hypothesis consists of several successive steps:

1. Develop the hypotheses. Explicitly formulate the null and the alternative hypotheses.
2. Specify the level of significance (5%, 1%).
3. Determine the direction of the test (one sided vs. two sided).

4. Select an appropriate statistical test (Z-test, t-test) and a sampling distribution (normal distribution, t-distribution, etc.).
5. Collect the sample data.
6. Estimate the value of the test statistic and/or the width of the confidence interval.
7. State the statistical conclusion.
8. Carry out the managerial decision.

Today, many consumer organizations perform checks of the truthfulness of information on a product package. Is the producer's claim regarding volume, fat percentage, nutritional content, and so on factually correct? (*see* OPPORTUNITIES AND CHALLENGES IN SOCIAL MARKETING).

Assume that a consumer organization suspects a producer of soft drink of underfilling its bottles. The package label states that a bottle contains 0.500 l.

- *Step 1*: First, two hypotheses are established:
  - H_0 = on average (as represented by the mean) a bottle contains 0.500 l or more
  - H_A = on average a bottle contains less than 0.500 l.
- *Step 2*: The level of significance is set to 0.05. In other words, the accepted error rate is 5%. If we conducted the test 100 times we would accept no more than 5 occurrences where H_0 is rejected given that it is, indeed, true. In the remaining 95 cases rejecting H_0 would be the correct decision because H_0 is false.
- *Step 3*: In the present case, the test is *unidirectional*. The consumer organization is only interested in testing whether the consumers get less soft drink than stated on the label. Provided that the producer, on average, puts more than 0.5 l in the soft drink, it is his problem. Providing more value than promised is not illegal.
- *Step 4*: The test to be used is a so-called Z-test according to the normal distribution (see below).
- *Step 5*: The consumer organization collects a sample of 30 soft-drink bottles from different retail outlets. Next, the content of each of the bottles is carefully measured. Table 1 shows the results. In the present case, the sample mean, $\bar{x}$, is 0.485 and the sample standard deviation, s, is 0.028. Since the sample size is ≥ 30 the central limit theorem applies, implying that the test can be carried out using the standard Z-test according to the normal distribution. Note that in statistical tests a sample is regarded "large" whenever $n \geq 30$.
- *Step 6*: The Z-test for a single mean is

$$Z = \frac{\bar{x} - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)} \tag{1}$$

In the present case, μ – the population mean – is set at 0.500, $\bar{x} = 0.485$, $n = 30$, and the population standard deviation σ is estimated by the sample standard deviation, $s = 0.028$. Thus,

$$Z = \frac{0.485 - 0.500}{\left(\frac{0.028}{\sqrt{30}}\right)} = -2.934 \tag{2}$$

- *Step 7*: To assess whether this value is statistically significant, one needs to inspect the standard normal distribution (see Figure 1). The threshold level for the 0.95 or 95% one-sided test is -1.645. The Greek letter α refers to the relative size of the area under the normal distribution *to the left of* -1.645. This implies that only in 5% of cases we would expect the z-value to be less than -1.645. Since our empirical value is -2.934, we reject H_0 and accept H_A implying that the mean bottle content is significantly less than 0.500 l (the actual probability value is 0.0016). The precise probability corresponds to somewhere between 1 and 2 cases out of 1000.
- *Step 8*: Conclusion: The consumer organization has statistical justification for taking legal action against the producer for underfilling its soft-drink bottles.

One- and Two-Tailed Tests

In the above example, the statistical test is carried out as a one-tailed or unidirectional test. The consumer organization's only concern is that the consumer gets no less than stated

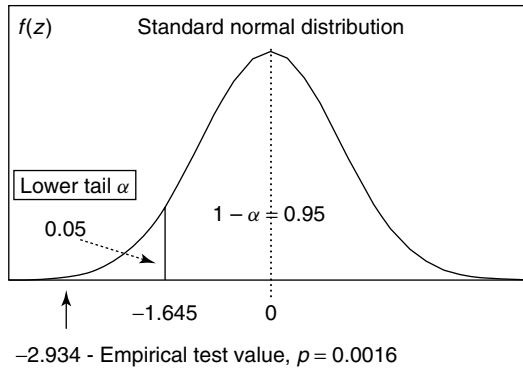


Figure 1 One-sided test.

Table 1 Exact content of 30 test bottles.

Bottle Number	*Liter*	*Bottle Number*	*Liter*	*Bottle Number*	*Liter*
01	0.492	11	0.498	21	0.475
02	0.487	12	0.497	22	0.526
03	0.479	13	0.477	23	0.503
04	0.491	14	0.481	24	0.447
05	0.519	15	0.507	25	0.450
06	0.476	16	0.499	26	0.474
07	0.480	17	0.499	27	0.499
08	0.468	18	0.442	28	0.546
09	0.532	19	0.448	29	0.469
10	0.491	20	0.418	30	0.476

on the label, that is, *at least* 0.500 l. A closer look at Table 1 reveals that in 6 out of the 30 cases the volume actually exceeds 0.500 l. From a producer's perspective, delivering more volume than promised may result in losses or at least in less profits than planned. Therefore, a decent management should strive for getting as close to the 0.500 l target value as possible. Delivering less than 0.500 l is regarded unfair to the consumer and may result in consumer complaints, while delivering more than 0.500 l is regarded as unfavorable to management since it decreases company profits.

In such a case, the statistical test is a two-tailed test. The computations are the same, but the implications from the test differ somewhat.

Assure that the above test value of 2.934 refers to a two-sided test (Figure 2 shows a two-sided test to be discussed below). In this case, we would reject H_0 either if (i) the observed z-value is less than -1.96 or (ii) if it is greater than $+1.96$. Otherwise, that is, if $-1.96 < z < +1.96$, we would accept H_0. Since the test is two sided, the probability of observing -2.934 is $2 \times (0.0016) = 0.0032$.

In the example of Table 1, the mean value of 0.485 is regarded *a point estimate*. It we assure that we are conducting a new test of $n = 30$, the mean would most probably not be 0.485 but rather 0.483, 0.491, or some other value.

Therefore, it is preferable to look at the *confidence interval* surrounding the point estimate. A 95% confidence interval represents a range. Given 100 measures we would expect the point estimate to fall within this span in 95 of the cases.

In the present case the confidence interval is

$$\bar{x} \pm z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) = 0.485 \pm 1.96\left(\frac{0.028}{\sqrt{30}}\right) = [0.475; 0.495] \qquad (3)$$

In other words, there is a 95% probability that the true but unknown population parameter of the mean will be somewhere between 0.475 and 0.495 l. It is noticed that even the upper bound (0.495) does not overlap with 0.500.

Suppose that we are very conservative, and that the test level of 5% error is regarded too liberal. Instead, we only accept a 0.1% level of error. In this case, the confidence interval

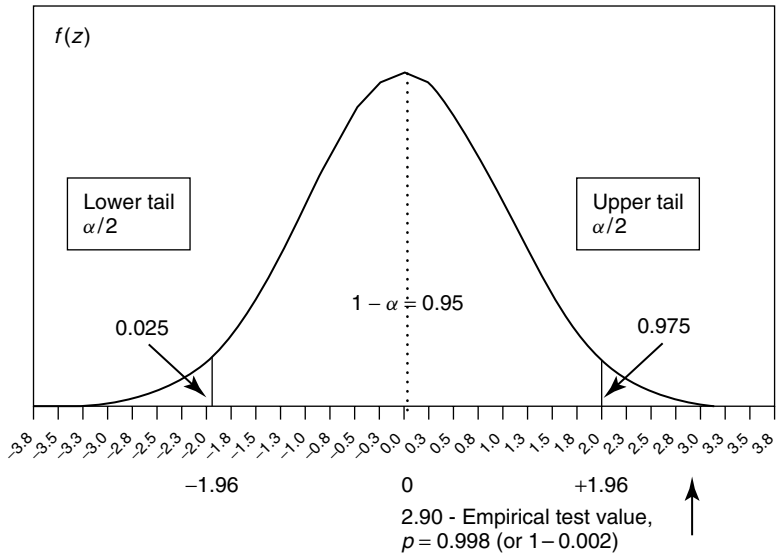


Figure 2 Two-sided test.

Table 2 Decisional dilemmas involved in hypothesis testing.

		State of Nature	
		H_0 True	*H_A True*
	Accept (fail to reject) H_0	Correct conclusion	Type II error
Decision	Reject H_0	Type I error	Correct conclusion

becomes

$$\bar{x} \pm z_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) = 0.485 \pm 3.09\left(\frac{0.028}{\sqrt{30}}\right) = [0.469; 0.501] \quad (4)$$

In this case, the upper bound of the confidence interval accurately exceeds 0.500. Therefore, if we want to be very sure about not falsely suing the company – that is, only accepting an error in 1 out of 1000 cases – we cannot rule out that the producer of soft drink is systematically *not* underfilling its bottles.

Notice that as we lower the probability of committing an error (increasing the error rate from 0.05 to 0.001) the width of the confidence interval increases.

Typical examples of two-sided tests are a candy bar or a soft drink containing a certain degree of sugar. If it contains too little sugar it tastes bitter and if it contains too much sugar it tastes too sweet. Likewise, most consumers prefer a specific amount of alcohol in red wine (13–14%). Also, in industrial production processes and when producing medical drugs well-specified two-sided tolerance levels are appropriate. Otherwise, the product quality suffers or the side effects increase beyond acceptable levels.

Type I and Type II Errors

Whenever testing a statistical hypothesis, one faces two decisional dilemmas: (i) the risk of rejecting something that is actually true (called a *type I error*) or (ii) the risk of accepting something that is not true or that is false (*type II error*) (see Table 2). Alpha (α) or the level of significance is the probability of committing a type I error, whereas the corresponding probability of committing a type II error is called

beta (β). There exists an inverse relationship between α and β. For a given sample size, β increases as α decreases. Stated differently, the more confident we want to be that our decision to reject H_0 is correct, the bigger is the chance that we accept a H_0 that is wrong. As we showed in the example with the soft-drink producer decreasing α from 0.05 to 0.001 implies that we fail to reject H_0. However, since our sample consisted of only 30 cases it is probable that new samples of $n = 30$ or a sample of $n = 300$ or 3000 will result in empirical means different from 0.485. Presuming that the soft drink example consist of 300 cases instead of 30, and supposing that the mean (0.485) and the standard deviation (0.028) is unchanged, the width of the 95% confidence interval narrows to [0.482; 0.488]. In this case, even the 99.9% confidence interval [0.480; 0.490] does not include 0.500, and thus we would definitely reject H_0 and recommend that legal action be taken against the producer for underfilling its bottles.

Generally, the probability of making a type II error increases as the value of the true population mean (estimated by the sample mean) is close to the H_0 value. Conversely, if the true population mean is far from H_0 the probability of committing a type II error is low.

Every day, we need to make decisions involving the risk of committing type I and type II errors. In the majority of cases, the consequences of the two error types are limited. Situations exist though, where making type I and type II errors are fatal. Imagine an airline pilot flying in darkness over mountains. Suddenly he gets a message from the Terrain Awareness and Warning System ("Terrain, Terrain") indicating that the plane is close to impact. At the same time, the plane's altimeter indicates that he is flying at an altitude of 6000 m. Which information should he believe? If he decides to take the warning seriously he would have to climb very quickly and pressure the airplane to its limits, thereby running the risk of its getting out of control and a part of the tail or even the fuselage breaking apart. On the other hand, if he should decide to ignore the warning system – and if it is correct – the plane will crash into the mountain within seconds.

Let us use the juridical system to illustrate the problem. In a criminal case, the accused is regarded as not guilty until the opposite is proved. In other words, H_0 = the accused is *not* guilty and H_A = the accused is guilty.

In this case, a type I error corresponds to a situation where the accused is innocent (H_0 – the attorney's claim is true) but where the person, nevertheless, is convicted and sent to prison (or even execution) for the crime. A type II error would correspond to the opposite situation: the accused is indeed guilty (H_A is true) but is released due to lack of evidence.

According to most lawyers, committing a type I error is worse than causing a type II error. In recent decades, the increasing use of DNA tests has helped in revealing and subsequently reversing a lot of type I and type II errors carried out by courts prior to the appearance of DNA tests.

Management also has to make decisions involving type I and type II errors. Consider a company that is about to launch a new product. Consequently, two hypotheses are established:

H_0 = the new product will not be profitable

H_A = the new product will be profitable.

Notice that we want to *reject* H_0, that is, we would like the new product to be successful.

If management makes the decision to launch the product (accept H_A) but the product turns out be a failure, (H_0 being true) a type I error has been committed. Usually, this is called *mismarketing*. Alternatively, if management decides not to launch (accept H_0) and it turns out that there is indeed a market (say, that a competitor markets a similar product that proves successful), this would be an example of a type II error. In most situations, a type I error (mismarketing) is regarded as more serious than a type II error (lost opportunity).

Sample Size and Test Statistics: Small Sample Test of the Mean

Up to now we have assumed that empirical tests were based on "large samples," implying that the sample size is at least 30 (the rationale of the minimum size of 30 is caused by the so-called *central limit theorem*). However, sometimes samples are small. In such cases, we must assume that the underlying population is normally distributed.

Suppose that in the example about underfilling of bottles, the available data consist of only the 10 observations in the second row of Table 1 (Bottles 1–10). In this situation, calculations are carried out in accordance with a t-test:

$$t = \frac{\overline{x} - \mu}{\left(\frac{s}{\sqrt{n}}\right)} = \frac{0.492 - 0.500}{\left(\frac{0.020}{\sqrt{10}}\right)} = -1.26 \quad (5)$$

The t-distribution resembles the standard normal distribution but differs somewhat in shape. And unlike the normal distribution, the t-distribution is characterized by a degree-of-freedom parameter. In the present case, the probability value associated to -1.26 with $n - 1 = 9$ degrees of freedom (assuming a two-tailed test) is 0.24 implying nonsignificance. Therefore, we cannot say that the company is underfilling its bottles. The same conclusion can be illustrated by computing the confidence interval around the mean, using a t-test with two tails and selecting an alpha level of 0.05.

$$\begin{aligned} \overline{x} \pm t_{\alpha/2} \times \frac{s}{\sqrt{n}} &= 0.492 \pm 2.28 \times \left(\frac{0.020}{\sqrt{10}}\right) \\ &= 0.492 \pm 0.014 \\ &\text{or } [0.478; 0.506] \end{aligned} \quad (6)$$

We notice that the upper bound of 0.506 overlaps the threshold level of 0.500. Thus we cannot be 95% sure that the company is underfilling its bottles.

Test of a Population Proportion: Large Sample

When dealing with a population proportion, computation of statistical significance and confidence intervals differ somewhat from when dealing with a population mean. Presume that the management of a retail chain specialized in fair trade clothes wants to find out if it has significantly more female customers than male customers. Management had decided that if the proportion of female customers is significantly bigger than the proportion of males, it will launch an advertising campaign in a regional magazine for women. Thus it conducts 400 interviews, with respondents randomly selected from the company's customer database. It turns out that 229 of the respondents are females.

Formally the hypotheses are H_0: $p \leq 0.50$ versus H_A: $p > 0.50$ (p is called the *population proportion*).

The appropriate test statistic is

$$z = \frac{\overline{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}} \quad (7)$$

It is a condition that $n \times p$ as well as $n \times (1 - \mathrm{p}) \geq 5$, where $\overline{p}$ is the sample proportion and p_0 is population proportion.

In the present case, $n \times p = 400 \times \left(\frac{229}{400}\right) = 229$ and $\mathrm{n} \times (1 - \mathrm{p}) = 400 \times \left(1 - \left(\frac{229}{400}\right)\right) = 171$. Since $229 > 5$ and $171 > 5$ we can use the z-test.

$$z = \frac{\left(\frac{229}{400}\right) - 0.50}{\sqrt{\frac{0.50(1 - 0.50)}{400}}} = \frac{0.0725}{0.025} = 2.90 \quad (8)$$

The related two-sided confidence interval is computed thus (assuming a 95% confidence interval and a two-tailed test):

$$\overline{p} \pm z_{.025}\sqrt{\frac{\overline{p} \times (1 - \overline{p})}{n - 1}} \quad (9)$$

Here

$$\begin{aligned} \left(\frac{229}{400}\right) \pm 1.96 \times \sqrt{\frac{\left(\frac{229}{400}\right) \times \left(1 - \left[\frac{229}{400}\right]\right)}{400 - 1}} \\ = 0.5725 \pm 0.0486 \\ \text{or } [0.524; 0.621] \end{aligned} \quad (10)$$

Since 2.90 exceeds 1.96 (and as the confidence interval does not include 0.5) we conclude that the proportion of female customers significantly exceeds that of male customers (the associated probability value of 2.90 is 0.998 (or $1 - 0.002$) (see Figure 2). On the basis of the test, we make the conclusion that H_0 is rejected: the retail store has significantly more female than male customers. Consequently, it is recommended that management advertises in magazines for women.

Table 3 Age of two samples of customers.

Downtown Supermarket				*Suburban Supermarket*				
28	41	43	42	41	36	42	35	54
44	35	44	58	36	38	46	37	23
39	41	45	48	28	36	22	38	
42	57	37	52	36	40	18	46	
20	45	34		57	35	38	29	
39	49	53		39	13	39	16	
34	45	41		14	29	40	35	
37	41	46		50	41	34	35	
38	43	34		40	35	39	37	
37	21	22		33	43	44	36	

Test of a Population Proportion: Small Sample

Tests for small samples concerning a population proportion are rarely used since the sample distribution follows a binomial distribution and therefore a normal approximation strictly does not apply. Several tests have been suggested, though. One such test is known as *Wilson's confidence interval* (Henderson and Meyer, 2001). Assume that we have conducted a small taste test among 20 persons ($n = 20$) and that 6 of them prefer the taste of our product to that of the competitor's product ($k = 6$). The $100(1-\alpha)\%$ Wilson confidence interval is

$$\frac{k + \frac{z_{\alpha/2}^2}{2}}{n + z_{\alpha/2}^2} \pm z_{\alpha/2} \times \sqrt{\frac{n \times \overline{p} \times (1-\overline{p}) + \frac{z_{\alpha/2}^2}{4}}{(n + z_{\alpha/2}^2)^2}} \tag{11}$$

In the present example – assuming a 95% two-sided test and that $\overline{p}$ is $6/20 = 0.3$ – we have

$$\frac{6 + \frac{(1.96)^2}{2}}{20 + (1.96)^2} \pm 1.96 \times \sqrt{\frac{20 \times 0.30 \times (1-0.70) + \frac{(1.96)^2}{4}}{(20+1.96)^2}} = 0.33 \pm 0.20 \text{ or } [0.13; 0.53] \tag{12}$$

Notice that 0.33 differs somewhat from $\overline{p}$ (0.30). Crow (1956) has developed tables of 90, 95, and 99% population proportion confidence intervals for all values of k for $n = 1$–30 using a Lagrangian interpolation. With $k = 6$, $n = 20$, and a 95% interval, the width of the interval becomes [0.14; 0.53], which comes very close to the Wilson estimate. Other publications addressing the small sample population proportion problem are Elston and Forthofer (1977) and Westfall (1985).

Test of Difference between Population Means: Two Large Samples

Up to now, all tests have been based on one sample where estimates of means and proportions have been tested against a hypothesized value (H_0). However, sometimes we are interested in comparing differences in means and proportions emanating from two independent samples. In the following, we show how this is done.

We begin with a test of difference between means assuming two big independent samples.

A retail chain is operating two supermarkets in the same city. One of the supermarkets is situated downtown, whereas the other one is in a suburban shopping center. Management is interested in figuring out whether the age of the two supermarkets' customers differs. If there is a significant difference in mean age between the customer bases of the two supermarkets, management might consider varying the supply of certain products between the two stores. So far the supply has been identical. The management makes a questionnaire and interviews a sample of customers at each supermarket (one of the questions concerns age). Table 3 shows the dataset and Table 4 displays the sample statistics.

We notice that the mean age of the downtown sample is higher than that of the suburban

Table 4 Sample statistics for Table 2 data.

Store	*Sample Size*	*Mean Age*	*Standard Deviation*
Downtown	$n_1 = 34$	$\overline{x_1} = 40.44$	$S_1 = 8.93$
Suburban	$n_2 = 42$	$\overline{x_2} = 35.79$	$S_2 = 9.54$

sample. But is the difference statistically significant? This can be estimated either by way of a z-test or by way of an interval estimate between the two means. The following test assumes that both samples are of size $n \geq 30$ and that σ_1 and σ_2 can be estimated by s_1 and s_2.

First, let us look at the z-statistic

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{13}$$

In the present case, we have

$$z = \frac{40.44 - 35.79}{\sqrt{\frac{8.93^2}{34} + \frac{9.54^2}{42}}} = 2.19 \tag{14}$$

The related probability value is 0.014 implying that we reject the hypothesis that there is no age difference between downtown and suburban customers.

The corresponding interval is estimated as

$$\bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2} \times \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \tag{15}$$

where $1 - \alpha$ is the confidence coefficient.

Assuming an α test level of 0.05 the interval becomes

$$\begin{aligned} 40.44 - 35.79 \pm 1.96 \times \sqrt{\frac{8.93^2}{34} + \frac{9.54^2}{42}} \\ = 4.65 \pm 4.16 \text{ or } [0.49; 8.81] \end{aligned} \tag{16}$$

We conclude that at a 95% level of confidence the margin of error is 4.16 years and the interval estimate of the difference between the mean ages of the customers of the two supermarkets is 0.49–8.81 years. On the basis of the dataset, it can be concluded that downtown customers are somewhat older than suburban customers.

However, the two samples are not of substantial size and thus very sensitive to small changes in the data. For instance, if the oldest person (age 58) in the downtown sample had instead been shopping at the suburban supermarket while the youngest person (age 13) in the suburban sample had visited the downtown store, then the confidence interval would change from [0.49; 8.81] to [−2.05; 6.55]. In that case, the lower bound of the confidence interval becomes negative implying that we cannot rule out (on the 95% level) that the mean age of the downtown shoppers is indeed *higher* than that of the suburban shoppers. We also notice that four of the suburban shoppers are teenagers. Is this a small group of school children who coincidentally dropped around and only bought one candy bar or a coke each? Most probably they are atypical shoppers. However, removing them from the computations would totally change the computations (the interval would be [−1.27; 6.25]). To sum up, it is strongly recommended that management collects more interviews so that it can base its decision on statistical estimates that are less sensitive to a few *influential observations* or *outliers*.

Test of Difference Between Means: Two Small Samples

In most cases, empirical samples conducted by or for companies consist of several hundred or even several thousand cases. However, as mentioned earlier, it may happen that the universe is limited. Many industrial companies have only a few hundred or so customers and, nevertheless, they may be interested in analyzing their customers' purchase behavior. If the number of cases sampled is less than 30 we cannot conduct a test like the one illustrated in the previous section.

First, we make the assumption that both populations have a normal distribution. Second, we presume that the population variances are equal, that is, $\sigma^2 = \sigma_1^2 = \sigma_2^2$. Third, since the variances normally are unknown, we assume that s, the standard deviation, can serve as a valid substitute for σ.

This being the case we can establish a pooled variance estimator:

$$s_{\text{pooled}}^2 = \frac{(n_1 - 1) \times s_1^2 + (n_2 - 1) \times s_2^2}{n_1 + n_2 - 2} \tag{17}$$

the t-test with $n_1 + n_2 - 2$ degrees of freedom becomes

$$t_{\alpha/2} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_{\text{pooled}} \times \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \tag{18}$$

Table 5 Sales (in 000 EUR) to customers.

Germany	*France*
982	842
1205	905
1031	1044
975	732
833	746
921	955
1114	1031
1082	958
1227	1089
1086	887
997	
701	

The corresponding confidence interval is

$$\overline{x}_1 - \overline{x}_2 \pm t_{\alpha/2} \times s_{\overline{x}_1 - \overline{x}_2} \tag{19}$$

Table 5 displays sales to a sample of an original equipment manufacturer company's industrial customers in Germany and France and Table 6 shows the sample statistics. Suppose that management would like to assess if the difference in mean between German and French sales is statistically significant.

First, we estimate the pooled variance estimator:

$$\begin{aligned} s^2_{\text{pooled}} &= \frac{(12-1)\times 148.98^2 + (10-1)\times 121.08^2}{12+10-2} \\ &= 18,804 \end{aligned} \tag{20}$$

Next, the difference in mean is tested by way of a t-test:

$$\begin{aligned} t_{\alpha/2} &= \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{s^2_{\text{pooled}} \times \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}; \\ &\textit{here}: \frac{1012.83 - 918.90}{\sqrt{18,804 \times \left(\frac{1}{12} + \frac{1}{10}\right)}} \\ &= 1.600 \end{aligned} \tag{21}$$

Table 6 Sample statistics for Table 5 data.

Country	*Sample Size*	*Mean Sale*	*Standard Deviation*
Germany	$n_1 = 12$	$\overline{x_1} = 1012.83$	$S_1 = 148.98$
France	$n_2 = 10$	$\overline{x_2} = 918.90$	$S_2 = 121.08$

The associated probability value with $n_1 + n_2 - 2$ degrees of freedom is 0.125 – a value that is not statistically significant.

The corresponding confidence interval (a t-test with 20 degrees of freedom and assuming a 95% confidence level) becomes

$$\begin{aligned} &1012.83 - 918.90 \pm 2.09 \\ &\quad \times \sqrt{18,804 \times \left(\frac{1}{12} + \frac{1}{10}\right)} = 93.93 \\ &\quad \pm 122.71 \text{ or } [-28.78; 216.64] \end{aligned} \tag{22}$$

Although the mean sale to German customers exceeds that to French customers, the difference is not statistically significant. Since the lower bound of the confidence interval is negative, we cannot rule out that the sale to French customers is actually greater than that to Germans.

Test of Difference between Population Proportions: Two Large Samples

In certain situations, management may want to compare two population proportions. We assume that the two samples are independent and that $n_1 \times p_1$, $n_1 \times (1 - p_1)$, $n_2 \times p_2$, and $n_2 \times (1 - p_2)$ are all ≥ 5. If these conditions are fulfilled, we can compute a z-value.

However, to begin with, we need two alternating hypotheses:

$H_0 : p_1 - p_2 = 0$ or

$H_1 : p_1 - p_2 \neq 0$.

First, assume that

$$\overline{p} = \frac{\overline{p}_1 + \overline{p}_2}{n_1 + n_2} \tag{23}$$

If this applies, the appropriate test statistics becomes

$$z = \frac{\overline{p}_1 - \overline{p}_2}{\sqrt{\overline{p} \times (1 - \overline{p}) \times \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \tag{24}$$

The interval estimate for the difference between the population proportions is computed thus:

$$(\bar{p}_1 - \bar{p}_2) \pm z_{\alpha/2} \times \sqrt{\frac{\bar{p}_1 \times (1-\bar{p}_1)}{n_1 - 1} + \frac{\bar{p}_2 \times (1-\bar{p}_2)}{n_2 - 1}} \quad (25)$$

Suppose that a new hair shampoo has been launched in two regions in opposite areas of a country. The two regions are regarded similar with respect to size of population and to sociodemographic characteristics. During a six-week period, the product is heavily advertised and the same amount is spent on the campaign in both markets. However, in the upper region all advertising money is spent on TV, whereas the money in the lower region is spent on a mix of TV, radio, newspaper, magazines, and Internet (Facebook, etc.).

Immediately after the end of the campaign, management hires a market research agency to carry out 1000 representative interviews in both regions asking respondents if they have seen and/or heard anything about the campaign. The results are as follows:

Upper region: $629/1000 = 0.629$

Lower region: $796/1000 = 0.796$

$$\bar{p} = \frac{629 + 796}{1000 + 1000} = 0.713 \quad (26)$$

$$z_{0.025} = \frac{0.629 - 0.796}{\sqrt{0.713 \times 0.287 \times \left(\frac{1}{1000} + \frac{1}{1000}\right)}} = -8.255 \quad (27)$$

The probability value of this estimate is less than 0.000. Therefore we reject H_0.

Selecting a 95% interval we get

$$(0.629 - 0.796) \pm 1.96 \times \sqrt{\frac{(0.629 \times 0.371)}{1000 - 1} + \frac{(0.796 \times 0.204)}{1000 - 1}} = -0.167 \pm 0.039 \text{ or } [-0.206; -0.128] \quad (28)$$

The interval implies that we can be 95% confident that the proportion of residents in the upper region who have become aware of the product is between 13 and 21% less than the proportion of residents in the lower region. Thus we have considerable evidence that the multimedia strategy that was being used in the lower region has been more successful with regard to generating attention than the single-media campaign of the upper region.

Test of Difference between Two Population Proportions: Two Small Samples

In the previous section, it was assumed that the two samples were big. Comparing two proportions where one or both samples are smaller than 30 rarely makes sense.

Nevertheless, a formula for the confidence interval has been developed for such situations. Agresti and Caffo (2000) suggest the following crude normal approximation:

$$(\bar{p}_1 - \bar{p}_2) \pm z_{\alpha/2} \times \sqrt{\frac{\bar{p}_1 \times (1-\bar{p}_1)}{n_1 + 3} + \frac{\bar{p}_2 \times (1-\bar{p}_2)}{n_2 + 3}} \quad (29)$$

According to Agresti and Caffo (2000) this test appears to perform quite well in many situations.

Assume that we have been interviewing 25 persons in the upper region out of which 16 have heard of the new product while 12 out of 15 in the lower region report knowledge of the new product. In this case, we have

$$(0.640 - 0.800) \pm 1.96 \times \sqrt{\frac{(0.640 \times 0.360)}{25 + 3} + \frac{(0.800 \times 0.200)}{15 + 3}} = -0.160 \pm 0.131 \text{ or } [-0.291; -0.029] \quad (30)$$

We notice that the confidence interval based on the small sample is longer than that of the big sample. However, the lower-region campaign is found to be more successful than the upper-region campaign.

Test of Difference between Two Means: Matched Samples

Up to now we have assumed that the two samples are unrelated or independent. Sometimes, it is of interest to take repeated measurements from the same individuals. In a blind test, a person may be asked to taste several soft drinks with different flavors and afterward the person is asked which one is preferred by way of a rating scale.

Another example of a related sample is a "before and after measurement." Presume that a sample of frequent customers of a company is asked to rate one of its well-known brands on a 0–100 temperature scale (the higher the better) according to purchase intention. Afterward they are exposed to a 2-min advertisement of the brand and subsequently they are asked to rate the brand once again. Table 7 shows an example of such a before-and-after rating.

The appropriate test is

$$t = \frac{\overline{d}}{\sqrt{\frac{1}{n \times (n-1)} \times \sum_{i=1}^{n} (d_i - \overline{d})^2}}; \text{ here we have:}$$

$$\frac{3.85}{\sqrt{\frac{1}{20 \times 19} \times 354.55}} = 3.986 \qquad (31)$$

The test value is t-distributed with $n - 1$ degrees of freedom. In the example, the 95% two-sided test with 19 degrees of freedom has a p-value of 0.00079. In other words, the probability that the two related samples are equal is less than 1 : 1000. Clearly, exposing persons to the advertisement has increased the brand's preference rating.

Since $t_{0.05/2;19} = 2.093$ it follows that the related confidence interval is

$$\overline{d} \pm t_{\alpha/2;n-1} \times \frac{s_d}{\sqrt{n}}; \text{ here } 3.85 \pm 2.093 \times \frac{4.32}{\sqrt{20}} \text{ or } [1.83; 5.88] \qquad (32)$$

We notice that the lower bound of the 95% confidence interval does not include 0.00. So, on the 95% level we reject the possibility that being exposed to the advertisement has lowered the purchase intention.

Table 7 Two related samples: before and after measurement.

Person	*Before*	*After*	*Difference*	$(Diff_i - Diff_{mean})^2$
1	65	72	7	9.92
2	59	60	1	8.12
3	78	78	0	14.82
4	67	66	−1	23.52
5	65	68	3	0.72
6	76	77	1	8.12
7	61	72	11	51.12
8	86	86	0	14.82
9	74	72	−2	34.22
10	88	97	9	26.52
11	68	73	5	1.32
12	64	62	−2	34.22
13	68	66	−2	34.22
14	82	92	10	37.82
15	58	68	10	37.82
16	80	87	7	9.92
17	72	77	5	1.32
18	65	70	5	1.32
19	84	88	4	0.02
20	73	79	6	4.62
Mean	71.65	75.50		Sum 354.55
		$Diff_{mean}$	3.85	
		$Diff_{std}$	4.32	

Testing for the Equality of Two Population Variances

Sometimes we are interested in differences in the *variance* of samples rather than in differences of means.

Assume a producer who has two production facilities for fabrication of red wine. Typically, a red wine contains around 12.5% of alcohol. The authorities allow for a 1.5% leeway in accuracy of the alcohol level. Government authorities regularly visit producers for taking samples of the alcohol content. If the alcohol content in a sample exceeds 14%, the wine cannot be sold in retail stores or to the catering sector. If the content is lower than 12.5%, the wine can be sold. However, in all probability, wine experts will give the wine unfavorable reviews and sales will decline.

Table 8 Alcohol content of wine.

Facility 1	*Facility 2*
13.29	12.16
11.64	12.41
13.79	12.45
12.44	12.27
12.61	12.45
12.82	12.62
11.75	12.44
12.30	12.88
11.97	12.05
12.22	12.46
12.46	
13.33	
$\overline{x_1} = 12.55, S_1^2 = 0.430$	
$\overline{x_2} = 12.42, S_2^2 = 0.054$	

Therefore management considers it very important to produce a wine that gets as close as possible to the 12.5% target. Thus it collects a sample of bottles from the two production facilities and gets them analyzed by a laboratory. Table 8 displays the results.

The appropriate test for testing the equality of variances is

$$F = \frac{s_1^2}{s_2^2}; \text{ here } F = \frac{0.430}{0.054} = 8.0 \qquad (33)$$

The degrees of freedom are f_1 (numerator) $= n_1 - 1 = 12 - 1 = 11$, and f_2 (denominator) $= n_2 - 1 = 10 - 1 = 9$.

The probability value of an F-value with 11 and 9 degrees of freedom is 0.002. Clearly, the variance of the wine percentage differs. There seems to be little doubt that the variance in alcohol percentage varies more in facility 1 than in facility 2, in spite of the fact that the mean percentage is nearly the same. It is recommended that management aims at providing a more stable production of wine in facility 1.

Hint: In the F-ratio always put the biggest variance in the nominator (as was done above). Otherwise, the test value would become $1/8 = 0.125$. The F-distribution is nonnegative and asymmetric implying that its peak usually is found left to the "center" (in such a case, the distribution is said to be *right skewed*). The F-test has two tails and, consequently, with a value of 0.125, the test would need to be carried out as a lower-tailed test meaning that the *smaller* the test value the smaller the probability. This often causes confusion even for people familiar with statistics. Readers who would like to know more about the F-test and about statistical distributions, in general, are referred to textbooks on business statistics like Anderson, Sweeny, and Williams (2003), Blackwell and Eldredge (2002) or Bowerman and O'Connell (2003).

Test of Difference between Two Means when Variances are Not Equal

When computing tests of significance and confidence intervals for the difference between means of two independent samples, it is usually assumed that the two population variances are equal. In such a case, the pooled standard deviation can be used as an estimator of the variance. However, as we have shown in the previous section, variances (as estimated by way of standard deviations) are not always equal.

If variances are not equal and if the sample size is small, the usual test cannot be conducted and a revised test must be employed.

Table 9 displays a small two-sample case where variances are obviously unequal.

Table 9 Measurements of cheese fat percentage.

	Facility A	*Facility B*
	18	24
	23	25
	28	26
	22	25
	25	
	17	
	16	
	19	
Mean	21	25
s^2	17.71	0.67

In this case, the ordinary t-test is replaced by a revised test statistic:

$$t' = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n2}}} \tag{34}$$

According to Satterthwaite (1946), t' can be estimated by way of an ordinary t-distribution if an approximate number of degrees of freedom is computable.

Assume that

$$v_1 = \frac{s_1^2}{n_1} \text{ and that } v_2 = \frac{s_2^2}{n_2} \tag{35}$$

When comparing the mean of the two samples, we presume that $w_1 = n_1 - 1$ and $w_2 = n_2 - 1$. In this case, the approximate number of degrees of freedom of t' is given by

$$v' = \frac{(v_1 + v_2)^2}{\frac{v_1^2}{w_1} + \frac{v_2^2}{w_2}} \tag{36}$$

then v' is to be rounded *down* to the nearest integer when using the t-distribution. Here we have

$$t' = \frac{21 - 25}{\sqrt{\frac{17.71}{8} + \frac{0.67}{4}}} = 2.592$$

and

$$v' = \frac{\left(\frac{17.71}{8} + \frac{0.67}{4}\right)^2}{\left(\frac{\left(\frac{17.71}{8}\right)^2}{7} + \frac{\left(\frac{0.67}{4}\right)^2}{3}\right)} = 7.99 \tag{37}$$

Since the probability value associated with $t = 2.592$ and 7 degrees of freedom (not 8!) is 0.0358, we reject the hypothesis (H_0) that the fat percentage of the two production facilities is the same. The fat percentage at facility B is higher but the variability across samples is more stable.

Notice that an ordinary t-test based on the Table 9 data with a pooled s^2 computed by employing formulas (17) and (18) above yeilds a value of -1.84 with an erroneous 10 degrees of freedom corresponding to a (two-tailed) probability value of 0.10. Thus, using an ordinary t-test we would erroneously accept H_0. The ordinary t-test tends to give too few significant results when the larger sample has the larger variance (as in our example) and too many when the larger sample has the smaller variance (see Snedecor and Cochran, 1989, pp. 96–99).

Miscellaneous

Hypothesis tests, probability values, and confidence intervals are sensitive to sample size. Therefore the researcher will often need to decide upon an appropriate *sample size* (*see* STATISTICAL APPROACHES TO DETERMINING SAMPLE SIZES) prior to gathering data. Another problem not addressed here concerns the ratio between size of the sample and of the population. If the population is less than 20 times bigger than the sample (if $n/N > 0.05$), confidence intervals for the mean and the proportion need to be adjusted or multiplied by the *finite population correction factor*, $\sqrt{\frac{N-n}{N}}$.

Notice that as n increases compared to N the correction factor decreases implying that the width of the confidence interval shrinks. Obviously, if $n = N$, that is, if the whole population is sampled (as census offices in some countries do in certain cases), then the correction factor becomes 0 and so does the width of the confidence interval. In such a case, there is no sampling error. Suppose that a company has 500 customers in its database and wants to survey 100. Since $500/100 = 5$, which is less than 20, the confidence intervals must be multiplied by $\sqrt{\frac{500-100}{500}} = 0.894$ (see Bowerman and O'Connell, 2003, pp. 283–286).

All statistical tests discussed in this section involve one or two samples. If one wants to test for the equality of means and more than two independent samples are involved, then the appropriate test is an *analysis of variance or ANOVA test* (*see* ANALYSIS OF VARIANCE AND COVARIANCE).

In case the assumptions underlying parametric tests are violated (homogeneity of variance/covariance, normality, etc.) researchers

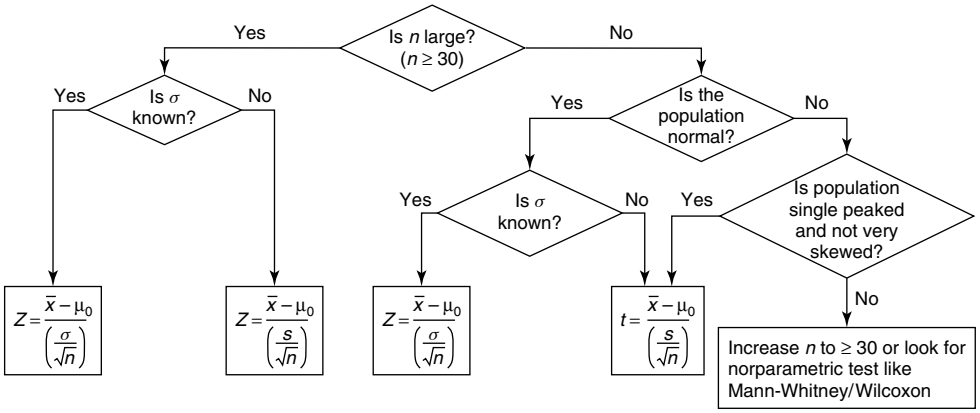


Figure 3 Selection of statistic for testing hypothesis about a population mean.

have developed an array of NONPARAMETRIC TEST (see also Conover, 1999).

We have provided tests for difference in means and proportions for large as well as for small samples above. With regard to small samples, an alternative to the suggested tests is known as *bootstrap statistics* (Mooney and Duval, 1993). Bootstrap statistics is a relatively new computer-intensive, general-purpose approach to statistical inference, falling within a broader class of resampling methods.

Bootstrapping involves taking a number of resamples of the observed dataset each of which is obtained by random sampling with replacement from the original dataset. Bootstrapping may also be employed for constructing hypothesis tests. It is often used as an alternative to inference based on parametric assumptions when assumptions are doubtful, or where parametric inference is not possible or requires complicated formulas for the calculation of standard errors. An alternative to Bootstrapping is known as *jackknife statistics* (Shao and Dongsheng, 1996).

Figure 3 provides a decisional diagram for selecting the appropriate statistical test for the mean.

ACKNOWLEDGMENTS

The author would like to thank Niels J. Blunch, Ole Drengsgaard, and Axel Schultz Nielsen for helpful comments on this article.

Bibliography

Agresti, A. and Caffo, B. (2000) simple and effective confidence intervals for proportions and differences of proportions – results from adding two successes and two failures. *The American Statistician*, **54**, 280–288.

Anderson, D.R., Sweeny, D.J., and Williams, T.A. (2003) *Modern Business Statistics Using Microsoft Excel*, South Western, Cincinnati.

Blackwell, K. and Eldredge, D.L. (2002) *Business and Economic Statistics Using Microsoft Excel*, South Western, Cincinnati.

Bowerman, B.L. and O'Connell, R.T. (2003) *Business Statistics in Practice*, McGraw-Hill, Boston.

Conover, W.J. (1999) *Practical Nonparametric Statistics*, John Wiley & Sons, Inc., New York.

Crow, E.L. (1956) Confidence intervals for a proportion. *Biometrika*, **43**, 423–435.

Elston, R.C. and Forthofer, R. (1977) Testing for Hardy-Weinberg Equilibrium in small samples. *Biometrics*, **33**, 536–542.

Henderson, M. and Meyer, M.C. (2001) Exploring the confidence interval for binomial parameter in a first course in statistical computing. *The American Statistician*, **55**, 337–344.

Mooney, C.Z. and Duval, R.D. (1993) *Bootstrapping: A Nonparametric Approach to Statistical Inference*, Sage Publications, Newbury Park.

Pan, W.Y. and Pooi, A.H. (2007) Testing the equality of means of two populations with unequal variances. *Malaysian Journal of Science*, **26**, 105–109.

Salsburg, D. (2001) *The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century*, Henry Holt & Company, New York.

Satterthwaite, F.E. (1946) An approximate distribution of variance components. *Biometrics Bulletin*, **2** (6), 110–114.

Shao, J. and Dongsheng, T. (1996) *The Jackknife and the Bootstrap*, Springer, New York.

Snedecor, G.W. and Cochran, W.G. (1989) *Statistical Methods*, Iowa State University Press, Ames.

Westfall, P. (1985) Simultaneous small-sample multivariate Bernoulli confidence intervals. *Biometrics*, **41**, 1001–1013.

latent class and finite mixture models

William R. Dillon

Introduction

Latent class analysis and more generally finite mixture models have seen increased use in marketing since the early 1980s. The popularity of latent class and finite mixture models can, in large measure, be traced to the important role individual differences (i.e., heterogeneity) play in understanding marketing and consumer behavior phenomenon. At the most basic level, latent class and finite mixture models have allowed marketing researchers to account for individual heterogeneity when estimating model parameters. Our objectives in this article are to provide an overview of latent class and finite mixture models and to provide a brief discussion of the various model forms and application settings that have relied on latent class and finite mixture models to account for individual-level heterogeneity.

Basic Model Forms

Latent class and finite mixture models refer to a class of model forms that recognize that not all individuals (of a population) are strictly alike. The term "latent" is used since the heterogeneity cannot be directly observed and the terms "class" or "finite mixture" are used to because the population is assumed to consist of homogeneous subgroups (or discrete classes), often referred to as the *components of the mixture*. We begin with a discussion of the more general finite mixture model and then turn our attention to the latent class model. A comprehensive treatment of many of the latent class and finite mixture models discussed herein, as well as others, can be found in Wedel and Kamakura (2000). Readers interested in more details on mixture distributions can refer to Titterington, Smith and Makov (1985).

Finite mixture models. Letting $f_s(y|\theta_s)$ be the probability density for observation Y, when sampled from the sth subgroup and α_s the fraction of the population that belongs to subgroup s, where $\alpha_1 + \alpha_2 + \cdots + \alpha_s = 1$, then Y has the mixture density of the form

$$f(y|P) = \sum_{s=1}^{S} \alpha_s f_s(y|\theta_s) \tag{1}$$

In Equation 1 P denotes the collection of unknown parameters $(\alpha_s, \theta_s), s = 1, 2, \ldots, S$. Put simply, the α's reveal the size of each of the subgroups, whereas the θ_s parameters reveal the distinctive character of each subgroup. Each individual in the population is assumed to belong to one of the S subgroups, denoted by $C_1, C_2, \ldots, C_s$, where subgroup membership for individual n is indicated by a $(S \times 1)$ vector z, with $z_s = 1$ if $y \in C_s$ and $z_s = 0$ otherwise; z is unobserved and is generally assumed to have a multinomial distribution, that is, $f(z|\alpha) = \prod_s \alpha_s^{z_s}$.

The conditional densities $f_s(y|P_s)$ can assume a variety of different multivariate distributional forms. For example, $f_s(y|P_s)$ can represent the multivariate normal or the multinomial distributions. When distributions are assumed to be discrete, for example, binomial, multinomial, or Poisson, finite mixture models are generally referred to as *latent class models*.

Parameters of the finite mixture model can be obtained via maximum likelihood (ML) estimation. For mixture distributions, the likelihood function assumes the following general form:

$$L(Y, P) = \prod_{i=1}^{N} \left(\sum_{s=1}^{S} \alpha_s f_s(y_i, \theta_s) \right) \tag{2}$$

By differentiating Equation 2 with respect to the elements of P, one can obtain the corresponding normal equations. It turns out that the ML estimates of α_s, *the mixing proportion* associated with subgroup s, is given by the sample mean of the posterior probability that y_i comes from that subgroup. Similarly, ML estimates of the θ_s parameters can be derived in the case where these parameters have explicit solutions by taking a weighted average of the ML equations, where the weights are the posterior probabilities.

ML estimates of the finite mixture model parameters have been obtained either by applying standard numerical optimization methods or the expectation maximization (EM)

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

algorithm (Dempster, Laird and Rubin, 1977). In the first approach, numerical optimization techniques, for example, the Newton–Raphson, simplex, or Fisher's scoring, are applied to the finite mixture model normal equations directly. In the EM approach unknown parameters are seeded with initial values and then the algorithm cycles through an *Expectation* (*E*) step and a *Maximization* (*M*) step. In the E-step the "missing data" (i.e., the subgroup that each individual belongs to) are estimated using the initial set of parameter start values and these "estimates" are then used to compute the posterior probabilities of subgroup membership. In the M-step, these "estimated" posterior probabilities are then used to obtain revised values of α's and θ's. The process cycles back to the E-step and continues in a similar fashion until the change in likelihood function values from one cycle to the next is judged to be small.

In empirical work, the number of subgroups S is not known and has to be inferred. The overwhelming temptation is to use the standard likelihood ratio test statistic, which is formed by taking 2 times the difference in the log-likelihood values for the s versus $s+1$ subgroup solutions. Typically, the likelihood ratio test statistic computed in this way has a chi-square distribution; however, in the case of mixtures, the test statistic is not asymptotically chi-square since the s subgroup model is at the boundary of the parameter space for the $s+1$ subgroup model. Because of this, a number of information statistics have seen use, all of which have the general form $-2\ln(L)+Qd$, where Q is the number of parameters and d is a penalty taking on different forms for different statistics. The oldest of the information statistics is Akaike's AIC in which $d=2$. Two other information statistics that have been used are the Schwarz's BIC, which sets $d=\ln(N)$, and Bozdogan's CAIC, which sets $d=\ln(N)+1$. The number of subgroups to retain is given by the model that produces the minimum value of the chosen information statistic. BIC and CAIC offer the advantage of being dimension consistent; that is, the probability of the statistic pointing to the correct number of subgroups converges to one as the sample size increases.

There are also a number of other bootstrap-type approaches. For example, Dillon and Mulani (1984) use a parametric bootstrap testing approach in which $T-1$ random Monte Carlo samples of size N are drawn from a population having s subgroups with conditional density $f_s(y|P_s)$. A finite mixture model is then fit with s and $s+1$ subgroups to each of the generated samples and the likelihood ratio statistic is computed. If the value of the likelihood ratio statistic obtained from the observed data exceeds $T(1-p)$ of the likelihood values obtained in the Monte Carlo samples, where p is the desired significance level, then the null hypothesis that the models with s and $s+1$ subgroups fit equally well cannot be accepted. Since this test and the three information statistics just described are based on the likelihood ratio statistic, all of these approaches should be viewed as heuristics.

Another important element in both latent class and finite mixture models are the *posterior probabilities*. Posterior probabilities give the probability that an individual belongs to each of the S subgroups and play an important role in the estimation process and in profiling the subgroups. The posterior probabilities are computed using Bayes rule and the estimated finite mixture parameters $\left(\hat{\alpha}_s, \hat{\theta}_s\right)$:

$$\hat{\pi}_s = f\left(z_s|y;\hat{P}\right) = \frac{\hat{\alpha}_s f_s\left(y|\hat{\theta}_s\right)}{\sum_{s'} \hat{\alpha}_{s'} f_{s'}\left(y|\hat{\theta}_s\right)} \qquad (3)$$

In empirical work, the posterior probabilities are often of particular interest as they are used as "new variables" in subsequent analyses to profile the segments and to explain subgroup membership as well as to assess the "quality" of the solution by forming an entropy-type measure.

Concomitant finite mixture models. Concomitant finite mixture models come about when one makes a distinction between variables that are chosen to form the basis on which the subgroups are identified (i.e., take on meaning) and another set of variables that are used to describe the subgroups and subgroup membership. We continue to denote the basis or core set of variables by y and now denote the concomitant or profile variables by x. There are two general types of concomitant finite mixture models. The first class of models include the core and descriptor variables simultaneously

within the finite mixture model and therefore attempt to identify the subgroups and, at the same time, discriminate across the subgroups, akin to discriminant analysis. Under this model form we can rewrite Equation 1 as

$$f(y, x|P) = \sum_{s=1}^{S} \alpha_s f_s(y|\theta_s) f_s(x|\phi_s) \qquad (4)$$

Here, conditional independence of the y and x variables given s is explicitly assumed, and, consequently, there is no need to denote these two variable types, other than for clarity of exposition.

In contrast, the second general model form treats the two set of variables very differently and uses the set of concomitant variables to simultaneously describe the subgroups. Although many variations of this approach have been proposed, one popular variant is to parameterize the prior probabilities as functions of the concomitant variables using the logit function:

$$\pi_{s|x} = \frac{\exp\left(\gamma_{0s} + x'\gamma_s\right)}{\sum_{s'} \exp\left(\gamma_{0s'} + x'\gamma_{s'}\right)} \qquad (5)$$

In Equation 5, γ_{0s} and γ_s are set to zero for identification. The specification of this finite mixture concomitant variable model now assumes the following form:

$$f(y|x; P) = \sum_{s} \pi_{s|x} f_s(y|\theta_s) \qquad (6)$$

where $P = (\gamma_0, \gamma_s, \theta_s)$. A comprehensive treatment of latent concomitant variable model forms can be found in Wedel (2002).

Latent class models. Letting $\mathbf{y}_i$ denote the response vector for the ith individual, where $\mathbf{y}_{ij}$ are polytomous responses, $r = 1, \ldots, R_j$, for $j = 1, \ldots, \mathcal{J}$ random categorical variables, and $c \varepsilon C, s = 1, \ldots, S$, denote the S latent classes, a general latent class model is defined by

$$f(\mathbf{Y}_i, c_s) = \prod_{j=1}^{\mathcal{J}} \prod_{r=1}^{R_j} \left[\theta_{jrs}\right]^{\delta_{ij}} \qquad (7)$$

and

$$f(\mathbf{Y}_i, P) = \sum_{s=1}^{S} \alpha_s f(\mathbf{Y}_i, c_s) \qquad (8)$$

where δ_{ij} is the Kronecker delta, which assumes a value of one if $y_{ij} = r$ and zero otherwise. Note that in addition to the usual constraint that, $\sum \alpha_s = 1$, it must also hold that $\sum_{j=1}^{R_j} \theta_{jrs} = 1$ for all j, s. The similarity of the latent class model defined by Equations 7 and 8 along with the appropriate constraints, and the finite mixture model should be obvious. However, there is one important difference, namely, the property that within a latent class the categorical variables are assumed to be statistically independent. This property, which is referred to as the *axiom of local independence*, means that within a latent class the joint densities can be expressed as a product of independent marginal densities. Though the property of local independence is used in other finite mixture models, it is not necessarily invoked in all varieties of finite mixture models. Both gradient methods and the EM algorithm have seen use in estimating the parameters (α_s's and θ_{jrs}'s) of the latent class model, although, since the ML equations for the basic latent class model have closed-form solutions, the EM algorithm is particularly well suited for this class of models.

The flexibility of the basic latent class of model can be greatly enhanced by adopting a logistic reparameterization of the basic latent class parameters (α_s, θ_s) in terms of real-value auxiliary parameters (ω_s, z_{jrs}), respectively, where these auxiliary parameters are themselves expressed as functions of a set of covariates. One set of covariates is presumed to influence the latent class sizes and the other set is presumed to influence the latent class conditional probabilities. Formann (1992) refers to this model form as the *linear logistic latent class model*.

Model Variants: Applications and Problem Settings

In this section, we provide a brief discussion of a number of latent class and finite mixtures model variants that have been proposed to address a wide variety of applications and problem settings.

Scaling models. Latent class scaling models have long history, dating back to the 1970s (cf. Goodman, 1975). These models impose a specific presumed scale structure on a set of

dichotomous or polytomous responses. These model forms, which are made possible because of the ability to impose constraints on the latent class parameters, are of two general types. *Purely deterministic scaling models* assume the absence of measurement error. Perhaps the best example is Guttman scaling wherein items are presumed to form a hierarchy of difficulty of agreeing with each of the k items and latent classes are introduced, which are isomorphic with each scale type, by imposing restrictions on the θ_{jrs} parameters. In contrast, *probabilistic scaling models* explicitly recognize and incorporate response error into the model. Goodman (1975) develops a number of Guttman-like models that can accommodate response errors by either incorporating an "intrinsically unscalable class" or a constant error rate parameter that governs the expected frequency of response error within or across latent classes; in addition, there have been several latent class Rasch and IRT models proposed as well.

Agreement models. Latent agreement models attempt to investigate the extent to which two or more observers agree in their assessment (or rating) of some phenomena. Clogg (1981) has applied latent class models in this context and argued that latent class agreement models have three salutary benefits related to their ability to (i) take into account chance agreement, (ii) afford coefficients that have clear interpretation, and (iii) identify sources of (dis)agreement.

Multidimensional scaling models. Latent class multidimensional scaling (LCMDS) has extended the applicability of conventional multidimensional scaling (MDS) models by accommodating samples of respondents that would otherwise be too large and messy for standard MDS approaches – even today conventional MDS models are geared to either individual-level data or small sample sizes. (Conventional MDS procedures are discussed in MULTIDIMENSIONAL SCALING OF PREFERENCE DATA.) Put simply, LCMDS perform multidimensional scaling while simultaneously clustering respondents – in marketing vernacular, LCMDS simultaneously estimates market segments as well as perceptual or preference structures of consumers (who belong to each of the latent segments). DeSarbo, Manrai and Manrai (1994) provide a comprehensive treatment of a variety of LCMDS models, including a generalized EM framework for the estimation of such models.

Regression models. Latent class regression models, or what are also referred to as *latent regression mixture models*, have been developed to explain the relationship between a set of independent, explanatory variables and a dependent variable when it is suspected that a single set of regression coefficients may be misleading because of respondent heterogeneity. (Multiple regression analysis is discussed in MULTIPLE REGRESSION) This class of models has its origins in the early work on "two-regime" switching regression models. A variety of latent class regression models have been developed to accommodate different distributional forms of the dependent variable. In the literature, we find latent class regression models for normal data (DeSarbo and Cron, 1988), for binary/multinomial data (Kamakura and Russell, 1989), for paired comparison data (Dillon, Kumar and de Borrero, 1993), and for count data (Wedel *et al.*, 1993). Mixtures of normals have seen extensive use in understanding preference and conjoint data, while mixtures of multinomials have been extensively used in the analysis of choice behavior. (Further details on conjoint analysis and logit models can be found in CONJOINT ANALYSIS and LOGIT MODEL, respectively.) A comprehensive treatment of the variety of latent class regression models that have been developed can be found in Wedel and DeSarbo (1994).

Structural equation models. Latent class structural equation models (SEMs) represent a fairly recent development (cf. Jedidi, Jagpal and DeSarbo 1997). (STRUCTURAL EQUATION MODELING in this volume of the encyclopedia discusses conventional SEMs.) Latent class SEMs allow the researcher to assess the measurement properties of a hypothesized theoretical model, that is, construct convergent and discriminant validity, disattenuate structural parameters between the exogenous "explanatory" constructs and the endogenous "dependent" constructs, while simultaneously

accounting for respondent heterogeneity. This class of models is analogous to fitting a multigroup structural equation model in the case where the groups are not directly known a priori. Latent class SEMs subsume a number of special cases. For example, if there are no endogenous constructs in the model, the model reduces to a finite mixture confirmatory factor model; if one assumes no measurement error, the model reduces to a mixture of simultaneous regressions or multivariate regressions.

Transition models. Latent class transition models represent a type of autoregressive modeling of longitudinal data. These model forms, which include hidden Markov models, consider several variables measured at each of several time points with a view toward explaining changes in the state variable. The state variable at one time point influences the state variable at the next time point in an autoregressive fashion. In these models two or more latent classes are introduced to distinguish between the state variable over time and, in so doing, assess true change from measurement error. Poulsen (1990) was the first to introduce the basic latent Markov model form.

Growth mixture models. Latent class growth analysis (LCGA) and growth mixture models (GMMs) are extension of conventional longitudinal growth analysis, which estimates mean growth curves that differ in intercepts and/or slopes. This analysis focuses on relating the repeated measurements of a dependent variable to a time-related variable (time scores), a time-varying covariate, and a time-invariant covariate, where the functional form of the relationship is parameterized in terms of random intercepts and slopes, which vary across individuals. In general, growth models can consider either continuous or categorical outcomes (dependent variables). LCGA and GMM extend these model forms by introducing latent classes that can either reflect homogeneous growth curves (i.e., members of a given latent class exhibit the same development) or capture the unique effects of a set of covariates on the within-class growth factors. Muthén (2004) provides a comprehensive summary of these model forms.

Bibliography

Clogg, C.C. (1981) New developments in latent structure analysis, in *Factor Analysis and Measurement in Sociological Research*, (D.J. Jackson and E.F. Borgatta), Sage, London, pp. 215–246.

Dempster, A.P., Laird, N.M. and Rubin, R.B. (1977) Maximum likelihood from the incomplete data via the EM-algorithm. *Journal of the Royal Statistical Society*, **B39**, 1–38.

DeSarbo, W.S. and Cron, W.L. (1988) A maximum likelihood methodology for clusterwise linear regression. *Journal of Classification*, **5**, 249–282.

DeSarbo, W.S., Manrai, A. and Manrai, L. (1994) Latent class multidimensional scaling: a review of recent developments in the marketing and psychometric literature, in *Handbook of Marketing Research*, (R. Bagozzi), Blackwell, Oxford, pp. 190–222.

Dillon, W.R., Kumar, A. and de Borrero, M.S. (1993) Capturing individual differences in paired comparisons: an extended BTL model incorporating descriptor variables. *Journal of Marketing Research*, **30**, 42–51.

Dillon, W.R. and Mulani, N. (1984) LADI: A latent discriminant model for analyzing marketing research data. *Journal of Marketing Research*, **26**, 438–458.

Formann, A.K. (1992) Linear logistic latent class analysis for polytomous data. *Journal of the American Statistical Association*, **87**, 476–486.

Goodman, L.A. (1975) A new model for scaling response patterns: an application of the quasi-independence concept. *Journal of the American Statistical Association*, **70**, 755–768.

Jedidi, K., Jagpal, H.S. and DeSarbo, W.S. (1997) Finite mixture structural equation models for response based segmentation and unobserved heterogeneity. *Marketing Science*, **16**, 39–59.

Kamakura, W.A. and Russell, G.J. (1989) A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research*, **26**, 379–390.

Muthén, B. (2004) Latent variable analysis: growth mixture modeling and related techniques for longitudinal data, in *The Sage Handbook of Quantitative Methodology for the Social Sciences*, (D. Kaplan), Sage, London, pp. 345–368.

Poulsen, C.S. (1990) Mixed markov and latent markov modeling applied to brand choice behavior. *International Journal of Research in Marketing*, **28**, 29–37.

Titterington, D.M., Smith, A.F. and Makov, U.E. (1985) *Statistical Analysis of Finite Mixture Distributions*, Wiley, New York.

Wedel, M. (2002) Conconcomitant variables in finite mixture models. *Statistica Neerlandica*, **56** (3), 362–375.

Wedel, M. and DeSarbo, W.S. (1994) A review of recent developments in latent class regression models, in *Handbook of Marketing Research*, (R. Bagozzi), Blackwell, Oxford, pp. 352–388.

Wedel, M., DeSarbo, W.S., Bult, J.R. and Ramaswamy, V. (1993) A latent class poisson regression model for heterogeneous count data. *Journal of Applied Econometrics*, 8, 397–411.

Wedel, M. and Kamakura, W.A. (2000) *Market Segmentation: Conceptual and Methodological Foundations*, Kluwer, Boston, MA.

models for categorical dependent variables

Gary J. Russell

Introduction

Researchers in marketing collect vast amounts of information on consumer behavior. Some of this information is continuous (such as total dollar expenditure for detergents in a given year) or takes the form of counts (such as the number of detergent bottles purchased in the last two months). However, much consumer purchase information is only available in the form of categorical information. For the purpose of this article, we define a *categorical dependent variable* as a nominally scaled variable that either classifies a consumer into a group or records consumer response to the marketing environment. For example, using a scanner panel dataset, a researcher might construct a categorical variable that reports the brand name and package size of a detergent purchase on a shopping trip. A model for a categorical dependent variable, then, is a predictive model that allows the researcher to make inferences about a categorical response, using additional information about consumers and the marketing environment.

We begin with a general discussion of the problems encountered in developing models for categorical variables. Although regression-based procedures are inappropriate in these settings, we argue that statistical tools known as *generalized linear models* (GLMs) provide a convenient approach for model construction. We then turn to issues of model specification and interpretation. To assist in this task, we briefly outline random utility theory (RUT) and show its relationship to GLMs. We conclude with a brief discussion of complex models addressing issues such as heterogeneity in response parameters and correlations among categorical variables in a repeated measures context.

Continuous Versus Discrete Variables

At the outset, it is useful to ask why researchers require distinct methodologies for dealing with categorical variables. Suppose, for example, the dependent variable y takes on two values, depending upon whether a consumer buys Brand A ($Y = 1$) or Brand B ($Y = 2$). By assembling the responses for all consumers into a vector $\mathbf{y}$, the researcher might analyze all observations in the dataset using regression software and obtain the coefficients β of the linear model $\mathbf{y} = X\beta$. Here, X is a matrix containing information on consumer demographics and marketing mix variables (such as price and promotion). Some of the X variables are continuous, but others could be categorical or ordinal.

The problem with the regression approach is not the variety of different types of X variables. Rather, the problem is the forecasts of the regression model. Because regression assumes that the dependent variable is continuous, forecasts of the dependent variable y may, in principle, be any number, positive or negative. For example, suppose that the forecast of the model for a particular set of X variables is 1.2. Does this mean that the consumer would buy Brand A? Does it mean that the consumer would buy some of Brand A and some of Brand B? Suppose, instead, that the forecast were 3.5. Does this mean that the consumer would buy Brand B? How should -1.7 be interpreted?

The problem shown here is one of logical consistency. Because the forecasts of the regression model cannot be constrained to take on only the values of the categorical dependent variable, the face validity of the regression approach is questionable. From a statistical perspective, hypothesis tests of the regression coefficients (such as t-tests) are also invalid because these tests are constructed on the assumption that the dependent variable is drawn from a normal distribution. Clearly, the normality assumption makes no sense for categorical variables. Taken together, lack of logical consistency and lack of statistical validity strongly argue that the categorical dependent variables should be modeled using specialized methods.

Generalized Linear Models

The key to modeling categorical dependent variables is to realize that, by definition, categorical dependent variables are discrete. Formally, if Y is a categorical dependent variable, then Y takes on one of R values for some known value R. We assume that, this variable is measured

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Commonly-used GLM models in marketing research.

Name of Model	*Type of Response Variable*	*Assumed Distribution of Response Variable*	*Mean of Response Variable E(Y)*	*Link Function g[E(Y)]*
Linear regression	Continuous	Normal	μ_h	$\mu_h \mathbf{x}(h)^{\mathrm{T}}\beta$
Poisson regression	Count	Poisson	λ_h	$\log(\lambda_h)\mathbf{x} = (h)^{\mathrm{T}}\beta$
Binary logit	Categorical ($R = 2$)	Binomial	$\pi(1)_h, \pi(2)_h$	$\log[\pi(1)_h/(1-\pi(1)_h)] = [\mathbf{x}(1,h) - \mathbf{x}(2,h)]^{\mathrm{T}}\beta$
Binary probit	Categorical ($R = 2$)	Binomial	$\pi(1)_h, \pi(2)_h$	$F^{-1}(\pi(1)_h) = [\mathbf{x}(1,h) - \mathbf{x}(2,h)]^{\mathrm{T}}\beta$, where F^{-1} (.) is the inverse of a N(0,1) cdf
RUT logit	Categorical ($R > 2$)	Multinomial	$\pi(1)_h, \ldots, \pi(R)_h$	$\log[\pi(k)_h/\pi(R)_h] = [\mathbf{x}(k,h) - \mathbf{x}(R,h)]^{\mathrm{T}}\beta$ for $k = 1, 2, \ldots, R-1$
Common logit	Categorical ($R > 2$)	Multinomial	$\pi(1)_h, \ldots, \pi(R)_h$	$\log[\pi(k)_h/\pi(R)_h] = \mathbf{x}(h)^{\mathrm{T}}\beta(k,R)\beta(k,R) = [\gamma(k) - \gamma(R)]$ for $k = 1, 2, \ldots, R-1$

on a nominal scale. That is, the numbers used to represent the R outcomes of Y convey no ordinal information, and certainly cannot be regarded as continuous measures. For example, in an analysis of the detergent market, the researcher might choose $R = 3$, and then assign the following values: $Y = 1$ (Tide), $Y = 2$ (Wisk), and $Y = 3$ (Cheer). In this article, we do not discuss how a categorical scale should be constructed. However, in general, the researcher must ensure that the classification covers all possible outcomes and that it uniquely assigns each outcome to one value of the Y variable.

The modeling strategy for a categorical dependent variable discards the idea that Y can be predicted directly. Instead, the researcher defines the probabilities $\pi(1), \pi(2), \ldots, \pi(R)$, where $\pi(j)$ is interpreted as the probability that Y takes on the value j. Because the $\pi(j)$ are probabilities, they must obey two logical constraints: all the $\pi(j)$ fall between the values 0 and 1, and the sum $\pi(1) + \cdots + \pi(R) = 1$. For example, when data analysis is conducted at the consumer level, we interpret $\pi(j)$ as the probability that the individual selects (or obtains) outcome j. The goal of the analysis, then, is to forecast the full set of probabilities that cover all the outcomes of Y. This implies that the researcher cannot predict with certainty which of the R outcomes will be actually observed. Putting this in statistical language, the strategy of the researcher is to predict the discrete probability distribution of Y, and not the actual realization of Y itself.

GLM modeling approach. One attractive way of implementing this strategy is to generalize regression analysis to deal with response variables that are not continuous. This is the approach taken by the GLM methodology (McCullagh and Nelder, 1989). When using the GLM approach, the researcher retains the idea that a linear predictor of the form $X\beta$ should be used to forecast the mean of the response variable. However, the distribution of the response variable need not be normal, and the link between the linear predictor and the mean response can be quite complex.

A summary of GLM models often found in marketing research is displayed in Table 1.

Briefly, the GLM methodology is built on the following key assumptions:

1. The response variable is some member of the exponential family of distributions. These include the normal, Poisson, binomial, and multinomial distributions, among others. As shown in Table 1, the binomial and multinomial distributions are appropriate for modeling categorical response variables.
2. The dataset contains $h = 1, 2, \ldots, N$ observations. We assume that the response variables are statistically independent across these N observations. This implies, in particular, that any dependence across observations must be built into the model by linking the means of the response variables across observations. Correlation structures across errors (such as serial correlations in a time series context) are not allowed.
3. The relationship between the linear predictor is given by a link function. Formally, a link function is some monotonic continuous function of the mean of the response variable of the form $g[\mathrm{E}(Y)]$, where $\mathrm{E}(Y)$ is the mean value of Y. In Table 1, we list the most common link functions used in practice. For a categorical response variable, researchers typically assume that the natural log of the ratio of outcome probabilities is equal to the linear function of the independent variables. Notice that, each link function allows the linear predictor to take on any value, positive or negative. For this reason, the estimated vector of parameters β is left unrestricted, a fact that simplifies model calibration. Nevertheless, the implied mean for the response variable will always obey the necessary logical constraints. For example, the binomial and multinomial models for categorical response variables generate a set of forecasted probabilities that lie between 0 and 1, and sum to 1 across all outcomes. Thus, by using an appropriate GLM, the researcher will obtain logically consistent forecasts.

GLM calibration and fit. Calibration of a GLM model is quite easy, because of readily available statistical software packages such as SAS and SPSS. The basic strategy is to write out the likelihood function of the GLM specification and then to maximize the likelihood with respect to the parameters β. As explained in McCullagh and Nelder (1989), a GLM model likelihood has a special structure that allows an iteratively weighted least squares routine to be used to generate parameter estimates. The important point for applied researchers is that, GLM models are numerically stable and converge quite rapidly. Standard errors of parameters are obtained via a transformation of the matrix of second derivatives of the log-likelihood function.

There is a strong temptation in GLM modeling to always interpret the β coefficients in the same manner as regression coefficients – even when the response variable is categorical. This is incorrect for two reasons. First, as noted above, the linear predictor $X\beta$ in a model with a categorical response variable does not yield a forecast of Y. Instead, it yields a highly nonlinear function of the mean of Y. Second, when analyzing categorical data, the researcher is actually fitting a discrete probability distribution to the R possible values of Y. Extreme care must be taken in both specifying the GLM model and in interpreting the results.

The well-known R^2 statistic for regression models cannot be used to assess the fit of a GLM model with a categorical dependent variable. Again, the problem is that R^2 is designed for continuous variables. An attractive alternative is the ρ^2 statistic (Louviere, Hensher, and Swait, 2000). The logic of the statistic relies on the observation that the likelihood of a categorical response GLM model is the product of probabilities across all observations – and so cannot exceed the value one. Accordingly, the log-likelihood (LogL) of any given model cannot be larger than zero. The procedure is to fit two models: one, a simple model with very few parameters, and the other, a more complex model with all the variables of interest. Then, using the LogL of the simple model (LogL(s)) and the LogL of the complex model (LogL(c)), we define

$$\rho^2 = \frac{[\mathrm{LogL}(s) - \mathrm{LogL}(c)]}{\mathrm{LogL}(s)} = 1 - \frac{\mathrm{LogL}(c)}{\mathrm{LogL}(s)} \tag{1}$$

This statistic lies between 0 and 1, with higher values indicating better fit. Extensive experience with real-world datasets indicates that models with excellent fit will have a ρ^2 value between 0.2 and 0.4 (Louviere, Hensher, and Swait, 2000). The choice of the simple model is somewhat controversial. However, most researchers choose a model that forces the discrete probability distribution of Y to be identical for all observations in the dataset.

Modeling Multinomial Responses

From the standpoint of applied research in marketing, the most difficult aspect in modeling categorical response data is developing an interesting way of predicting the probabilities of response $\pi(1)_h, \ldots, \pi(R)_h$. Although, the link functions shown in Table 1 are commonly used in marketing research, they are not the only choices available. We first discuss the properties of the models in Table 1, and then turn to a variety of alternatives.

Let h denote an observation and k denote one of the R categorical values of the response variable Y. We assume that for each h and k, the researcher has access to Q independent variables in the form of the vector

$$\mathbf{x}(k,h) = [\mathrm{x}(k,h,1), \mathrm{x}(k,h,2), \ldots, \mathrm{x}(k,h,Q)]^{\mathrm{T}} \tag{2}$$

At the outset, it is important to note that these variables can be either continuous variables or (dummy-coded) categorical variables. For example, if the goal of the model is to predict brand choice, h might denote a household and k a particular brand of detergent selected on a shopping trip (*see* PURCHASE, SCANNER, AND MEDIA PANELS). Then, we might define one variable of $\mathbf{x}(k,h)$ to be the price of brand k observed by household h during the purchase occasion. We could also define another component of $\mathbf{x}(k,h)$ to be a binary (0,1) indicator variable for which the value 1 denotes the presence of a promotion for brand k during the shopping trip. It is customary to include an intercept variable in each $\mathbf{x}(k,h)$ to represent the average level of $\pi(k)_h$ across all observations. These are $R-1$ binary (0–1) indicator variables of the form $\mathrm{x}(k,h,q) = 1$ only if $k = q$. For reasons that will become clear shortly, $\mathrm{x}(k,h,R)$, the intercept for the Rth outcome provides redundant information and so should not be included in the model.

The models for categorical response shown in Table 1 are known as *logit models* (*see* LOGIT MODEL). For the $R > 2$ case, researchers generally use the term *multinomial logit*. Consider the link function

$$\log\left[\frac{\pi(k)_h}{\pi(R)_h}\right] = [\mathbf{x}(k,h) - \mathbf{x}(R,h)]^T \beta \tag{3}$$

for $k = 1, 2, \ldots, R-1$. Using the fact that the probabilities $\pi(1)_h, \ldots, \pi(R)_h$ must sum to 1, we obtain the general expression

$$\pi(k)_h = \frac{\exp\left[\mathbf{x}(k,h)^{\mathrm{T}}\beta\right]}{\left\{\exp\left[\mathbf{x}(1,h)^T\beta\right] + \cdots + \exp\left[\mathbf{x}(R,h)^{\mathrm{T}}\beta\right]\right\}} \tag{4}$$

where exp[.] denotes the exponential function and the summation in the denominator runs over the R possible outcomes of the response variable. Equation 4 is the familiar "us versus (us + them)" formulation for probabilities and market shares often used in the marketing research literature.

Logit model identification. It must be emphasized that the link between independent variables and the probability of response is highly nonlinear. Accordingly, the researcher must take great care in choosing independent variables to include in the $\mathbf{x}(k,h)$ vectors. To understand the key issues, notice that the logit link function in Equation 3 is defined in terms of the differences between two vectors of variables $[\mathbf{x}(k,h) - \mathbf{x}(R,h)]^T\beta$. Because these link functions define the response outcome probability distribution, we must conclude that only the differences between the independent variables of the various outcomes impact model fit.

This can lead to problems in some applications. For example, suppose that h denotes household and that the researcher wishes to examine the impact of household income on the response variable. Denote income as $\mathrm{x}(k,h,q)$. (That is, income is the qth variable in the

vector $\mathbf{x}(k,h)$.) Because income depends upon household h, but not upon outcome k, it must be the case that $x(k,h,q) = x(k*,h,q)$ for any alternative k^* different from k. Put another way, $[x(k,h,q) - x(k*,h,q)] = 0$ because all these variables have the same value, namely, the income of the household. In statistical jargon, the coefficient on household income is said to be *unidentified* under these circumstances. If the researcher ignores this logical constraint and attempts to estimate the multinomial logit using GLM software, the algorithm will fail to converge because there exists no unique value of the response coefficient vector β that maximizes the likelihood function. Using a slightly more complex argument, it can also be shown that the Rth outcome intercept is unidentified in the presence of intercepts for outcomes $1, 2, \ldots, R-1$. An excellent discussion of identification issues in the context of a categorical response variable model may be found in chapter 2 of Train (2003). In general, independent variables are only useful if they impact the relative sizes of the outcome response probabilities.

Model identification issues can be addressed by using theory (from psychology or economics) to suggest how variables should enter the model. For example, Kalyanam and Putler (1997) use microeconomic theory to argue that household income should impact choice in the form of an *interaction* between price and income. (In this context, interaction means that the price of a particular outcome should be multiplied by household income.) Accordingly, these authors estimate a multinomial logit model including two price variables: brand price, and the interaction of brand price, and household income. They show empirically that higher income households are less price sensitive than lower income households. Clearly, some variables cannot be used directly in the $\mathbf{x}(k,h)$ vector, but can be introduced in a special way (using interactions) to modify the β coefficients of other variables (such as price).

RUT logit model. The major justification for Equation 4 is RUT. Briefly, the RUT paradigm assumes that consumer h assigns a utility $\mathrm{U}(k,h)$ to each of the possible alternatives according to the expression

$$U(k,h) = \mathbf{x}(k,h)^T \beta + \varepsilon(k,h) \tag{5}$$

where $\varepsilon(k,h)$ is a stochastic error term reflecting unobservable factors that influence choice. On each choice occasion, the consumer is assumed to select the product k that yields maximum utility. Because the error is unobserved, RUT requires the analyst to use the stochastic properties of the $\varepsilon(k,h)$ variables to derive the probability that alternative k will be chosen. If the $\varepsilon(k,h)$ are independent draws from a Gumbel (extreme value) distribution, then the choice probabilities $\pi(k)_h$ can be shown to be equal to the multinomial logit model of Equation 4.

Detailed discussions of RUT and the logit model can be found in Ben-Akiva and Lerman (1985), Louviere, Hensher, and Swait (2000), and Train (2003). The classical marketing application of the multinomial logit to scanner panel data is Guadagni and Little (1981). This article also provides excellent practical advice on the development of $\mathbf{x}(k,h)$ variables that capture consumer brand preference, choice process dynamics, and marketing mix effects.

Common logit model. Not all marketing research applications require a model that conforms to RUT. For example, consider the multinomial logit link function

$$\begin{aligned} \log[\pi(k)_h/\pi(\mathrm{R})_h] &= \mathbf{x}(h)^{\mathrm{T}}[\gamma(k) - \gamma(R)] \\ &= \mathbf{x}(h)^{\mathrm{T}}\boldsymbol{\beta}(k,R) \end{aligned} \tag{6}$$

for $k = 1, 2, \ldots, R-1$. Notice that, the vector of independent variables $\mathbf{x}(h)$ depends only on the observation h, not upon the alternative chosen. This implies the relationship

$$\pi(k)_h = \frac{\exp\left[\mathbf{x}(h)^{\mathrm{T}}\beta(k,R)\right]}{\left\{\begin{array}{c}\exp\left[\mathbf{x}(h)^{\mathrm{T}}\beta(k,R)\right] + \cdots \\ + \exp\left[\mathbf{x}(h)^{\mathrm{T}}\beta(R,R)\right]\end{array}\right\}} \tag{7}$$

where $\exp\left[\mathbf{x}(h)^{\mathrm{T}}\beta(R,R)\right] = 1$ because the logic of Equation 6 requires $\beta(R,R) = \mathbf{0}$. An alternative (and logically equivalent) expression for the

probabilities is

$$\pi(k)_h = \frac{\exp\left[\mathbf{x}(h)^{\mathrm{T}}\gamma(k)\right]}{\left\{\begin{array}{c}\exp\left[\mathbf{x}(h)^{\mathrm{T}}\gamma(l)\right]+\cdots\\ +\exp\left[\mathbf{x}(h)^{\mathrm{T}}\gamma(R)\right]\end{array}\right\}} \quad (8)$$

where we assume (without loss of generality) that $\gamma(\mathbf{1}) + \cdots + \gamma(R) = 0$. Using Equation 8, it is easy to show that the link function of this model can be also written as

$$\log\left[\frac{\pi(k)_h}{\phi_h}\right] = \mathbf{x}(h)^{\mathrm{T}}\gamma(k) \quad (9)$$

where ϕ_h is the geometric mean of the $\pi(k)_h$ probabilities. Again, we require that the sum of $\gamma(\boldsymbol{k})$ coefficients equals zero. Although, Equations 6–9 appear very different, they all describe the same model.

In the statistics literature, the relationship in Equation 8 is usually defined as a multinomial logit model. For clarity, we will refer to Equation 4 as the *RUT logit* model and Equation 8 as the *common logit* model. It should not be surprising that the specification issues of the common logit are very different from those of the RUT logit. The key property of the common logit is that the independent variables $\mathbf{x}(h)$ vary by observation h, but not by the level k of the outcome variable Y. In addition, the parameters of the common logit model (denoted $\gamma(k)$) vary by alternative. Recall, that the RUT logit specification has one common set of parameters (β)for the model, but a set of independent variables $\mathbf{x}(k,h)$ that vary by alternative. In a certain sense, the RUT logit and the common logit are mirror images of one another.

The common logit is often the only type of logit model that is supported by standard statistical software. The form of the model is motivated by the observation that

$$\begin{aligned}\log\left[\frac{\pi(k)_h}{\phi_h}\right] &= \log[\pi(k)_h] - \log[\phi_h]\\ &= log[\pi(k)_h] - \mathrm{E}_k[\log[\pi(k)_h]]\end{aligned} \quad (10)$$

where $E_k[\log[\pi(k)_h]]$ is the mean (over k outcomes) of the $log[\pi(k)_h]$ values. (This follows because ϕ_h is the geometric mean of the choice probabilities.) Accordingly, the Equation 9 link function requires that the linear predictor $\mathbf{x}(h)^T\gamma(k)$ be equal to the deviation of the log-probability of outcome k around the mean of the log-probabilities for all R outcomes.

This notion of modeling deviation scores using a linear model is very similar to the logic used by statisticians in developing ANOVA (analysis of variance) decompositions for factorial experiments (*see* ANALYSIS OF VARIANCE AND COVARIANCE). The fact that all the $\gamma(k)$ parameters sum to zero is fully consistent with standard ANOVA parameter-coding schemes. For this reason, if all the explanatory variables in $\mathbf{x}(h)$ are categorical, the researcher can directly use ANOVA coding to analyze a categorical outcome variable within the common logit framework. Moreover, it can be shown that the resulting model is actually a special case of a log-linear model. Log-linear models are Poisson regression models that allow the analysis of multiway contingency tables. Thus, there is an intimate connection between the common logit and contingency table analysis (*see* CROSS-TABULATION).

Interpretation of common logit parameters can be very difficult. For most applications, the researcher estimates the $\beta(k,R)$ coefficients (using Equation 7), and not the γ(k) parameters (using Equation 8). For example, consider a choice model developed by Jank and Kannan (2005) using a generalization of the common logit model. Each consumer in the study was offered three alternatives: (i) buying a printed copy of a research article, (ii) buying a PDF version of the same article, or (iii) not buying the article in any format. We display the $\beta(k,R)$ coefficients reported by Jank and Kannan (2005) on the left-hand side of Table 2. On the right-hand side of Table 2, we display the corresponding $\gamma(k)$ coefficients, computed from the $\beta(k,R)$ estimates by using Equation 6 and the fact that the $\gamma(k)$ sum to zero across alternatives. Note that, each row of the $\gamma(k)$ section of Table 2 adds to zero, as required.

The prices of the print copy and the PDF electronic file are the key variables. Two versions of the PDF price are used: PDF price as a percentage of the print price, and PDF price in absolute units. Consistent with Equation 7,

Table 2 Common logit parameter estimates.

	$\beta(k,R)$ *Coefficients*			$\gamma(k)$ *Coefficients*		
	Print	*PDF*	*None*	*Print*	*PDF*	*None*
Intercept	0.315	1.835	0	−0.402	1.118	−0.717
Print price	−0.015	0.013	0	−0.014	0.014	0.001
PDF price (percentage)	0.009	−0.065	0	0.027	−0.046	0.019
PDF price (actual)	0.097	0.038	0	0.052	−0.007	−0.045

The $\beta(k,R)$ coefficients are reported by Jank and Kannan (2005). The $\gamma(k)$ coefficients are inferred using the $\beta(k,R)$ values and the link function in Equation 6.

all variables enter into the expressions for all alternatives, and the column corresponding to the last alternative (nonpurchase, denoted by None) is set to zero. At a minimum, we would expect that the signs of the two versions of PDF price in the PDF column should both be negative (because higher PDF prices reduce the probability of buying the article in PDF format). This is not true in the PDF column of the $\beta(k,R)$ section of the table, but it is true in the PDF column of the $\gamma(k)$ section. This occurs because the $\beta(k,R)$ coefficients are measured relative to the Rth level of the response variable (in this case, None), whereas the γ(k) coefficients reflect deviations around the common geometric mean of all outcome probabilities.

Latent Variable Models

The RUT logit and common logit models are routinely used in marketing research applications. The models are easy to compute and often provide useful forecasts. Nevertheless, for some marketing applications, it is desirable to develop more complex models that take into account additional facts about the problem setting.

The conceptual tool that allows new models to be constructed is the notion of a latent variable. Simply put, a latent variable is a random variable whose outcomes cannot be directly observed. However, latent variables impact observed variables in a known way. In the context of the present discussion, a latent variable formulation allows the researcher to derive a new link between independent variables and the probability that the categorical outcome variable will take on certain values. This new link function, along with the assumption of a multinomial distribution, allows the GLM framework to generate a new method of analyzing categorical response data.

Ordinal regression. The simplest model of this sort is ordinal regression. The model assumes that the R states of the categorical outcome variable can be ordered in a natural way owing to the manner in which the data were collected. For example, survey respondents might be given a five-point Likert scale and asked to rate a product in terms of preference. The researcher assumes that for each person h, there exists a latent preference variable

$$V(h) = \mathbf{x}(h)^{\mathrm{T}}\beta + \varepsilon(h) \tag{11}$$

where $\varepsilon(h)$ is a random draw from some known distribution with mean zero. In other words, V(h) is a random variable with mean $\mathbf{x}(h)^T\beta$.

The R points of the rating scale are assumed to correspond to $R-1$ cut points $(\tau_1, \ldots, \tau_{R-1})$ on the latent continuum of V(h). If V(h) falls below τ_1, then the respondent picks the lowest point on the rating scale ($Y = 1$). If V(h) falls between τ_1 and τ_2, then the respondent picks the next higher point on the rating scale ($Y = 2$). This logic continues for all R outcomes of Y. By making distributional assumptions about the error $\varepsilon(h)$, it is possible to develop a formula for $\pi(k)_h$, the probability that respondent h (with background characteristics $\mathbf{x}(h)$) will have a V(h) value that falls into the region of the latent continuum corresponding to $Y = k$. Using the GLM framework for multinomial distributions, the researcher is then able to obtain estimates of β and all the scale cut points $\tau_1, \ldots, \tau_{R-1}$. This

procedure is known as *ordinal logistic* or *ordinal probit*, depending upon whether the distribution of $\varepsilon(h)$ is assumed to be logistic or normal, respectively. A detailed discussion of the model can be found in Bock (1975).

Random utility theory models. It should be now apparent that the latent preference Equation 11 is very similar to Equation 5, the expression for random utility introduced earlier in the context of the RUT logit model. In consumer choice applications, the general approach is to gather the utilities of all k alternatives into one vector $\mathbf{U}(h)$ and then write

$$\mathbf{U}(h) = X\beta + \mathbf{E}(h) \tag{12}$$

where each row of X contains equal $\mathbf{x}(k,h)^T$ and the kth element of the $\mathbf{E}(h)$ vector contains the error term $\varepsilon(k,h)$. As before, the RUT paradigm assumes that the consumer chooses alternative k if the utility $\mathrm{u}(k,h)$ is the largest. Usually, we would write the model in Equation 12 in such a way that the mean of $\mathbf{E}(h)$ equals zero.

By assigning a specific multivariate distribution to the vector $\mathbf{E}(h)$, we can derive the probability that item k will have the maximum utility of all brands. This, of course, is $\pi(k)_h$, the probability that categorical outcome level k will be selected by consumer h. When the distribution of $\mathbf{E}(h)$ is a member of the generalized extreme value (GEV) family, the resulting probability expression is called a *GEV logit model* (Ben-Akiva and Lerman, 1985). Models of this sort are generalizations of the RUT logit discussed earlier. Alternatively, if $\mathbf{E}(h)$ **has** a multivariate normal distribution, then the probability expressions yield a *multinomial probit* model (Train, 2003). Both of these models can be viewed as different methods of linking the $X\beta$ linear predictor to the probability distribution of the categorical response variable. Thus, the GEV logit and the multinomial probit are also GLM models.

Computationally, GEV logit models are easier to estimate because the link functions have a known analytic form. In contrast, multinomial probit link functions are dependent upon an $R-1$ dimensional integral. For this reason, researchers can simply append additional code to a multinomial response GLM algorithm to estimate a GEV logit model. In contrast, estimation of multinomial probit parameters requires the use of simulation estimation technology to make the estimation of parameters practical. Train (2003) is an excellent source of information on simulation approaches to model calibration.

Both the GEV logit and the multinomial probit allow the researcher to alter the probability distribution of the outcome variable in a complex manner. Consider, for example, the nested logit model, a type of GEV logit that assumes a partitioning of the choice alternatives into subgroups. Let $\pi(s,j(s))_h$ denote the probability of selecting product $j(s)$ within the subgroups. The nested logit model decomposes this probability as

$$\pi(s,j(s))_h = \pi(s)_h \pi(j(s)|s)_h \tag{13}$$

where $\pi(s)_h$ is the probability of selecting subgroups, and $\pi(j(s)|s)_h$ is the probability of selecting product $j(s)$, given that subgroup s has been selected. In this model, $\pi(j(s)|s)_h$ takes the form of an RUT logit for brands within the submarket. The first term, $\pi(s)_h$, is also a logit model, but includes a specialized variable (called the *inclusive value*) that measures the relative attractiveness of each subgroup. Derivation of the nested logit model (and other GEV logit models) can be found in Ben-Akiva and Lerman (1985). Models with properties similar to the nested logit can also be constructed using the multinomial probit framework. See Train (2003) for details.

Parameter heterogeneity. Another important use of latent variables is the modeling of UNOBSERVED HETEROGENEITY in link function parameters. Researchers in marketing, generally, assume that consumer characteristics, (about which the researcher has limited knowledge) will induce variation in the β-portion of the GLM linear predictor. For example, suppose we rewrite the link function of Equation 3 as

$$\log\left[\frac{\pi(k)_h}{\pi(R)_h}\right] = [\mathbf{x}(k,h) - \mathbf{x}(R,h)]^{\mathrm{T}}\,\beta(h) \tag{14}$$

where the parameter vector $\beta(h)$ is now allowed to depend on h. Equation 14 then implies an RUT logit model for each observation h, but

with parameters that vary across h. To calibrate a model of this sort, $\beta(h)$ is treated as latent variable drawn from some known distribution. *Latent class models* (Kamakura and Russell, 1989; Wedel and Kamakura, 1997) assume that the probability distribution of $\beta(h)$ is discrete (*see* LATENT CLASS AND FINITE MIXTURE MODELS). An E–M (expectation–maximization) algorithm can then be used to calibrate the model. *Random coefficient models* assume that the distribution of $\beta(h)$ is continuous. Most random coefficient models are now calibrated using a hierarchical Bayes technology (Rossi, Allenby, and McCulloch, 2005) (*see* RANDOM COEFFICIENTS MODELING). In either case, specialized algorithms are required for a GLM model with unobserved parameter heterogeneity.

MULTIVARIATE CATEGORICAL DATA

One of the most challenging forecasting problems in marketing research is the modeling of multivariate categorical data. To fix ideas, define $Y(h,j)$ as a categorical variable for respondent h. Instead of just one Y, we assume the researcher has access to M different categorical variables ($j = 1, 2, \ldots, M$) arranged as a column vector

$$Y(h) = [Y(h,1), Y(h,2), \ldots, Y(h,M)]^{\mathrm{T}} \quad (15)$$

In the language of experimental design theory, the M variables $Y(h,j)$ are called *repeated measures* on respondent h. For example, in a survey analysis setting, index j could stand for different product names. In this context, the $Y(h,j)$ might be preference ratings for M different products on a Likert scale. Alternatively, in a panel data setting, index j could be defined as the time of the purchase occasion. Then, the $Y(h,j)$ would denote M different product choices (at M different times) by respondent h. The key question is how to adapt the univariate GLM framework discussed earlier to modeling the multivariate $\mathbf{Y}(h)$ vector.

Conditional independence. Three different solutions have appeared in the marketing literature. By far, the most common approach is to assume that the repeated measures of respondent h are *conditionally independent*. The basic idea is to assume that the parameter vector of the link function for each $Y(h,j)$ is $\beta(h)$, a set of parameters that do not depend on the particular categorical variable j being modeled. Moreover, given the $\beta(h)$vector for respondent h, the M outcome variables in $\mathbf{Y}(h)$ are assumed to be statistically independent. Conditional independence allows the researcher to treat repeated measures as independent univariate GLM models, subject to the restriction that all the link functions of respondent h have the same set of parameters. Correlations between the $Y(h,j)$ can be introduced into the model, but only by using the link functions to connect the $\pi(k)_h$ outcome probabilities of the various categorical variables. Guadagni and Little (1981) provide an interesting illustration of the conditional independence approach. They define the index j to be choice occasion (time) and allow the current choice probabilities to depend upon the past realizations of $Y(h,j)$.

Multivariate probit. The second approach involves generalizing the latent variable system introduced earlier in Equation 12. Suppose that we partition the $\mathbf{U}(h)$ utility vector as

$$\mathbf{U}(h) = [\mathbf{U}(h,1), \mathbf{U}(h,2), \ldots, \mathbf{U}(h,M)]^{\mathrm{T}} \quad (16)$$

where $\mathbf{U}(h,j)$ contains the set of utilities needed to predict the outcome of $Y(h,j)$. Consistent with RUT, we assume that the alternative with the highest value among the elements of $\mathbf{U}(h,j)$ will become the alternative reported by $Y(h,j)$. Aside from this partitioning of the $\mathbf{U}(h)$ vector, the basic RUT framework of Equation 12 is retained. If we assume that $\mathbf{E}(h)$ is drawn from a multivariate normal distribution, then the researcher can derive the probability that item k will be the outcome for $Y(h,j)$. However, because the errors in $\mathbf{E}(h)$ are correlated across the $Y(h,j)$ variables, the outcome probabilities of the $Y(h,j)$ categorical variables will be correlated as well. This approach is known in the marketing literature as the *multivariate probit* model. (It should not be confused with the multinomial probit model discussed earlier.) An application of the multivariate probit to market basket analysis is found in Manchanda, Ansari, and Gupta (1999).

Autologistic regression. The third approach is to view the $\mathbf{Y}(h)$ vector as a random draw from a multivariate distribution for M categorical variables. For example, suppose that each of the

$Y(h,j)$ are binary (0–1) indicator variables, and define $\tau(j)_h$ as the *conditional probability* that $Y(h,j)$ equals 1, given the known outcomes of all other variables $Y(h,r)$ for all r different from j. Further, define the link function

$$\log\left[\frac{\tau(j)_h}{(1-\tau(j)_h)}\right] = \mathbf{x}(h,j)^{\mathrm{T}}\beta + \sum_{i\neq j}\theta(i,j)Y(h,j) \quad (17)$$

where $\mathbf{x}(h,j)$ are explanatory variables and the $\theta(i,j)$ are symmetrical parameters ($\theta(i,j) = \theta(j,i)$) measuring the association among the $Y(h,j)$ outcomes. Equation 17 is called an autologistic regression *model* because the outcome of $Y(h,j)$ depends on all other variables in $\mathbf{Y}(h)$. It is possible to show that the M link functions corresponding to Equation 17 imply that the joint distribution of the $Y(h,j)$ variables is given by an M-dimensional *multivariate logistic* distribution. The amount of correlation between the $Y(h,j)$ is controlled by the $\theta(i,j)$ parameters. If all the $\theta(i,j)$ are equal to zero, then each of the $Y(h,j)$ variables is a statistically independent binary logit GLM. Russell and Petersen (2000) use the multivariate logistic model to analyze the impact of price within and across four paper goods categories. Generalizing this theory for multivariate R-level categorical variables remains a topic for future research.

Summary

The analysis of categorical dependent variables requires special care in both model specification and parameter interpretation. Researchers must understand that models for categorical data are designed to forecast the discrete probability distribution of categorical outcomes, not the specific outcome itself. By far, the most prominent analysis tools used in marketing research are the RUT logit and the common logit. These models can be viewed as special types of GLMs. More complex models can be developed through the use of latent variables. In particular, latent variables permit the construction of models that incorporate ordinal relationships among outcomes, differing patterns of similarity among alternatives, and heterogeneity in response parameters. Recent research has begun the development of multivariate models that allow for correlations among categorical dependent variables.

Bibliography

Ben-Akvia, M. and Lerman, S.R. (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge.

Bock, R.D. (1975) *Multivariate Statistical Methods in Behavioral Research*, McGraw-Hill, NewYork.

Guadagni, P.M. and Little, J.D.C. (1981) A logit model of brand choice calibrated on scanner data. *Marketing Science*, **2**, 203–237.

Jank, W. and Kannan, P.K. (2005) Understanding geographical markets of online firms using spatial models of customer choice. *Marketing Science*, **24**, 623–634.

Kalyanam, K. and Putler, D.S. (1997) Incorporating demographic variables in brand choice models: an indivisible alternatives framework. *Marketing Science*, **16**, 166–181.

Kamakura, W.A. and Russell, G.J. (1989) A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research*, **26**, 379–390.

Louviere, J.J., Hensher, D., and Swait, J. (2000) *Stated Choice Methods: Analysis and Application*, Cambridge University Press, Cambridge.

Manchanda, P., Ansari, A., and Gupta, S. (1999) The shopping basket: a model for multicategory purchase incidence decisions. *Marketing Science*, **18**, 95–114.

McCullagh, P. and Nelder, J.A. (1989) *Generalized Linear Models*, Chapman and Hall, London.

Rossi, P., Allenby, G.M., and McCulloch, R. (2005) *Bayesian Statistics and Marketing*, John Wiley & Sons, Inc., London.

Russell, G. and Petersen, A. (2000) Analysis of cross category dependence in market basket selection. *Journal of Retailing*, **76**, 367–392.

Train, K.E. (2003) *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge.

Wedel, M. and Kamakura, W.A. (1997) *Market Segmentation: Conceptual and Methodological Foundations*, Kluwer Academic Publishers, Dordrecht.

multidimensional scaling of preference data

Wayne S. DeSarbo and Crystal J. Scott

Introduction

Using the broad conceptualization by Carroll and Arabie (1980), we define multidimensional scaling (MDS) as a family of various geometric models for the multidimensional representation of the structure in data as well as the corresponding set of methods for fitting such spatial models. Carroll and Arabie (1980) present a taxonomy of the area of MDS based on the properties of the input measurement data (e.g., number of modes, number of ways, power of the mode, scale type, conditionality, completeness of the data, and replications) and properties of the underlying multidimensional measurement model (e.g., type of geometric model, number of sets of points in the derived space, number of derived spaces, and degrees of constraint on model parameters). Thus, their definition extends classical MDS, which typically deals only with spatial models for proximity data (e.g., similarities) to various other types of continuous and discrete representations, as well as to other data types. Our focus is upon two major types of models used for the analysis of two-way dominance data (i.e., preference, consideration to buy, or choice data) as is typically collected in marketing research: the vector MDS model and the unfolding MDS model. Readers interested in a more comprehensive discussion of this broad area of MDS are encouraged to read the very comprehensive book by Borg and Groenen (2005) for an in-depth treatment of these and other types of MDS approaches for the analysis of such data (e.g., correspondence analysis). For expositional purposes, we assume that the data to be analyzed is a two-way data set of metric brand preferences where the rows of this data matrix (**P**) reflect a sample of consumers and the columns of the matrix reflect brands in a designated product/service class. The general entry in this data matrix (P_{ij}) is therefore, the metric preference rating (*see* PRIMARY SCALES OF MEASUREMENT) given for brand j by consumer i. The objective of the MDS models to be described is to estimate a spatial configuration (a joint space) of both row (consumers) and column (brands) objects such that their geometric interrelationships most parsimoniously recover the input preference data **P**.

The Vector MDS Model

Tucker (1960) and Slater (1960) were the first to independently formulate this scalar products based model for geometrically displaying the structure in such two-way data (Carroll, 1980). Related to factor analysis (*see* EXPLORATORY FACTOR ANALYSIS), the underlying model can be mathematically represented as

$$P_{ij} = \sum_{t=1}^{T} a_{it} b_{jt} + e_{ij} \qquad (1)$$

where

$i = 1, \ldots, I$ consumers;
$j = 1, \ldots, J$ brands;
$t = 1, \ldots, T$ dimensions;
a_{it} = the tth coordinate of the terminus of the preference vector for consumer i in the derived space;
b_{jt} = the tth coordinate of the location of brand j in the derived space;
e_{ij} = error.

We explain this geometric representation via Figure 1 that illustrates the workings of this particular model in Equation (1) for the simple case of two dimensions, four brands, and three consumers. The two dimensions are labeled as such in the figure and represent typical scatter plot axes. The brand coordinates (b_{jt}) are plotted here for each of the four brands (A, B, C, and D) and represent the positions of the brands in this derived space. Note, the consumer locations (a_{it}) are represented in this model (labeled Consumers 1, 2, and 3) as vectors emanating through the origin whose orientations point in the direction of increasing preference or utility for each consumer. Each of the three consumers' vectors point in different directions reflecting heterogeneity in their respective tastes and preferences. (Note that, we draw the tails of the vectors as dashed lines reflecting the areas of the space that are not preferred for each consumer).

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

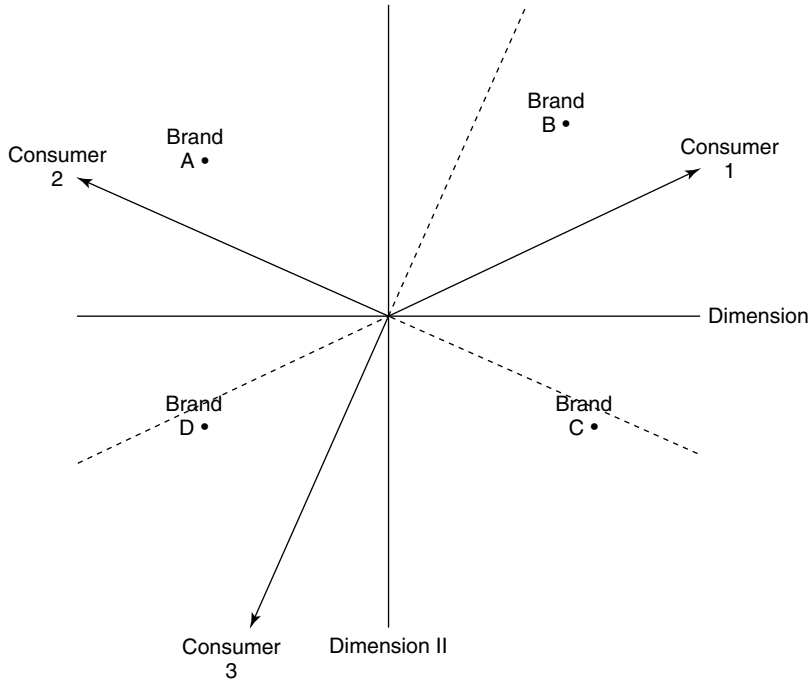


Figure 1 The vector model.

The predicted cardinal preference values are given by the orthogonal projection of each of the brands onto each consumer's preference vector. Thus, Consumer 1 has the following order of predicted preference: B, C, A, D; Consumer 2: A, D, B, C; and Consumer 3: D, C, A, B. Note, the consumer vectors are typically normalized to equal length in such joint space representations although, under certain data preprocessing conditions, the raw lengths of such vectors can be shown to be proportional to how well each consumer's preferences are recovered by the vector representation. The goal of such an analysis is to simultaneously estimate the vectors and brand coordinates in a given dimensionality that most parsimoniously captures/recovers the empirical preference data. The analysis is repeated for $t = 1, 2, 3, \ldots, T$ dimensions and the dimensionality is selected by inspection of t versus a goodness-of-fit statistic (e.g., variance accounted for) that measures how well the model predictions in Equation (1) match the input preference data, given the number of model parameters being estimated. Note, the cosines of the angles each consumer vector forms with the coordinate axes render information as to the importance of these derived dimensions to that consumer. The isopreference contours for this model in two dimensions for a particular consumer vector (i.e., locations of equal preference) are lines perpendicular to a consumer vector at any point on that vector, since brands located on such lines would project at the same point of predicted preference onto the vector. Thus, it is important to note that this vector model is *not* a distance-based model. Also, one can freely rotate the joint space of vectors and brand points and not change the model predictions (the orthogonal projections of the brand points onto the consumer vectors) or goodness-of-fit results. As noted by Carroll (1980), one of the unattractive features of this vector model is that it assumes that preference changes monotonically with respect to all dimensions. In other words, since a consumer's vector points in the direction of increasing preference or utility, the more of the dimensions in that direction imply greater preference; that

is, *the more the better*. In marketing, this can create conceptual problems depending upon the nature of the underlying dimensions. This assumption may not be realistic for many latent attributes or dimensions underlying brands in a product/service class. For example, it is not clear that consumers prefer infinite amounts of large size and sportiness (assuming those were the two underlying dimensions driving their preferences) in their family car. In addition, it would imply that the optimal positioning of new brands would be out to infinity, which is most often not realistic.

The Simple Unfolding Model

Coombs (1950) (unidimensional unfolding) and Bennett and Hays (1960) (multidimensional unfolding) introduced a different type of geometric MDS model for the analysis of such metric preference data called the *simple unfolding model*. Mathematically, the model can be represented as

$$F(P_{ij})^{-1} = \sqrt{\sum_{t=1}^{T}(a_{it} - b_{jt})^2} \qquad (2)$$

where we use the same notation as that used in the vector model in Equation (1) with some striking differences relating to the nature of how consumers' preferences are represented in this approach. In particular, in the simple unfolding model, both brands and consumers are represented as coordinate points in the derived T-dimensional space. Here, the closer the distance between a consumer and a particular brand, the higher the predicted preference for that particular brand. This, distance is inversely related to preference and that is why one sees an inverse function for preference on the left-hand side of Equation (2) above. Figure 2 illustrates this model in two dimensions with three consumers and four brands. The order of predicted preference

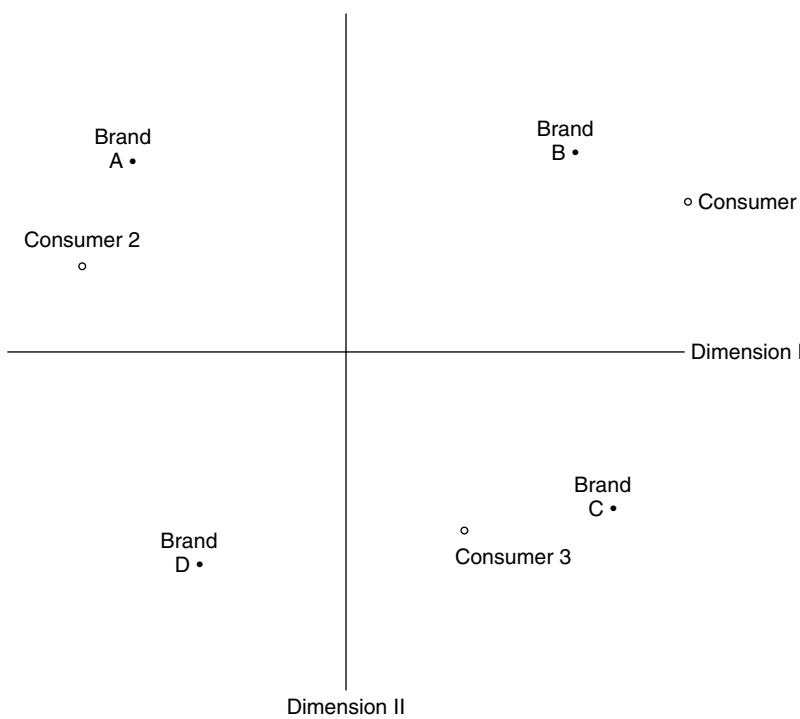


Figure 2 The simple unfolding model.

for Consumer 1 would be B, C, A, D; for Consumer 2 this would be A, D, B, C; and, for Consumer 3 would be C, D, B, A. This geometric model is Euclidean-distance based, and the consumer locations are often called *ideal points* as they reflect the optimal combination of the dimensions for that consumer. Thus, unlike the vector model where higher preference tends off to infinity in the direction of the estimated consumer vector, the simple unfolding model posits highest utility at the ideal point, and preference trails off uniformly in any direction from that ideal point. The further away a brand is from a consumer ideal point, the less preferred that brand is. Here, the isopreference contours in two dimensions are concentric circles around an ideal point since all brand points of fixed radius around the particular brand point would reflect equal preference by this model. Both the brand locations and consumer ideal points are estimated to best recover the empirical preference data. As in the vector model, the analysis is conducted in $t = 1, 2, 3, \ldots, T$ dimensions, and the dimensionality is selected in terms of parsimony in contrasting goodness of fit versus the number of model parameters estimated. Like the vector model, one can rotate the joint space configurations and not affect distances. The simple unfolding model also is indeterminate with respect to the origin of the derived joint space. Nonetheless, the simple unfolding model is more appealing to marketers given its finite preference utility assumptions and intuitive underlying Euclidean distance model. Ideal points often represent targets for marketers who try to attempt to position their brands near target segment ideal points (DeSarbo and Rao, 1986). Carroll (1980) also introduced the weighted and general unfolding models which provide for additional forms of individual differences in a multidimensional unfolding context, but these more complex models involve the estimation of many more parameters and successful applications of these more complex unfolding models in the marketing literature are lacking.

As first noted by Carroll (1980), while the vector and ideal point model appear to be quite different geometric representations, one can easily show that the vector model is a special case of the simple ideal point model. If one were to move an ideal point for any individual consumer further and further out along a fixed line from the origin while holding the brand locations fixed, one ends up with a vector utility model for that consumer in the limit where the ideal point tails off to infinity. The family of concentric circles surrounding that ideal point begins to flatten out and resemble the isopreference lines of the vector model. Thus, there is additional flexibility involved in the simple unfolding model which can accommodate the vector model as a special case. However, despite its greater flexibility and intuitive appeal, the simple unfolding model suffers from frequent degenerate, uninformative solutions where the brand points and consumer ideal points are estimated to be excessively separated from each other (see Borg and Groenen, (2005) for a discussion of attempts to resolve this difficulty).

A Marketing Application

In this section, we present a detailed marketing application of these two MDS models using customer consideration-to-buy data from a tracking study conducted in 2002 by a large US automotive consumer research supplier. The study has been conducted semiannually in June and December for over 20 years across several different vehicle segments. It is used to gauge automotive marketing awareness and shopping behavior among vehicle intenders. An *intender* is defined as a prospective buyer who will be "in market" or has plans to purchase a new vehicle within 6–12 months. The surveys used in these tracking studies were conducted among new vehicle intenders and were collected from an automotive consumer panel of more than 600 000 nationally represented households (see DeSarbo, Grewal, and Scott 2008 for additional details). Using a four-point preference response scale (4 – "definitely would consider", 3 – "probably would consider", 2 – "probably would not consider", and 1 – "definitely would not consider"), the respondents rated each brand in terms of their consideration to buy (preference) corresponding to the product segment in which they intend to purchase.

For this illustration, we used the large sports utility vehicle segment which includes

the following 16 brands: Chevy Suburban, Chevy Tahoe, Cadillac Escalade ESV, Cadillac Escalade EXT, Ford Expedition, Ford Excursion, GMC Yukon, GMC Yukon Denali, H1 Hummer, H2 Hummer, Lexus LX470, Lincoln Navigator, Mercedes G-Class, Range Rover, Toyota Land Cruiser, and Toyota Sequoia. Using this product class, we estimate both the vector model and the ideal point (unfolding) model. The data matrix for both models consists of the same individual stated consideration to buy from 278 respondents on the set of 16 large sports utility vehicles SUVs listed above. Thus, this is a two-way dataset of metric brand considerations/preferences where the rows of this data matrix are the 278 consumers and the columns represent the 16 brands.

The vector model results. We first consider the vector model where the goal is to simultaneously estimate the consumer vectors and brand coordinates in a given dimensionality while best recovering the empirical consideration data. Figure 3 shows the resulting configuration in two dimensions with brand locations represented as diamond-shaped points and the consumers as vectors emanating from the origin (we normalize the vectors to equal length for convenience and drop the tails of the vectors in an attempt to reduce the clutter in the figure). With respect to Dimension I (horizontal axis), we see that the GM vehicles are separated from the non-GM vehicles. We thus label this dimension as GM versus non-GM. Dimension II (vertical) represents price with the least expensive vehicles at the top of the axis and increasing in price as

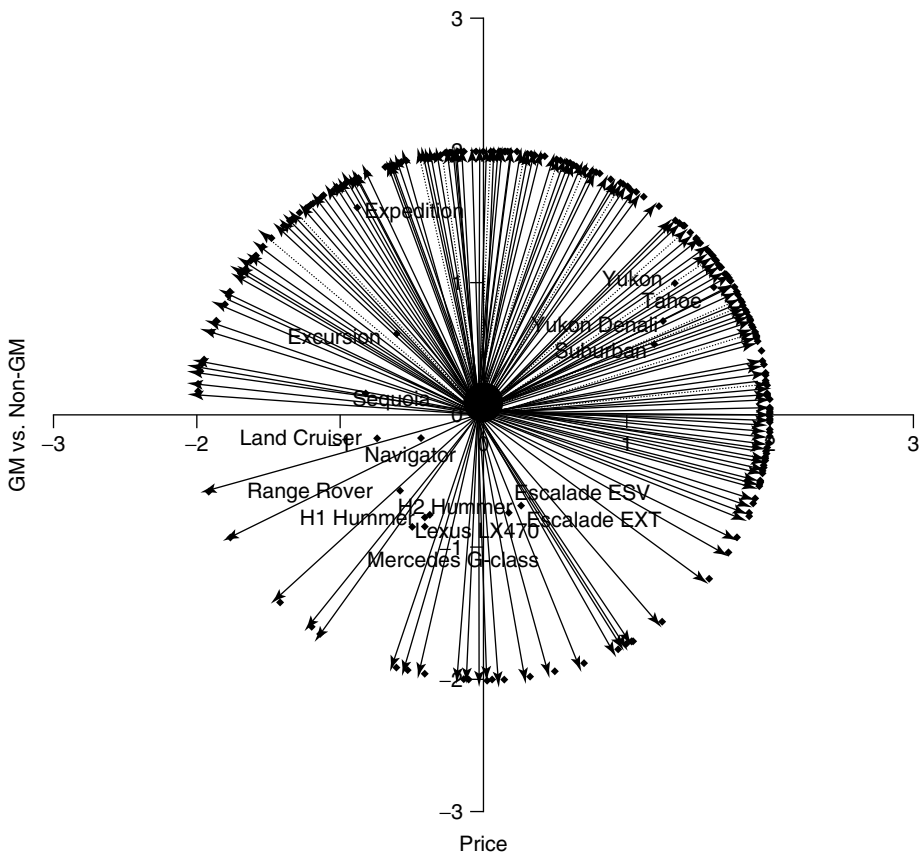


Figure 3 Vector model – SUV application.

you travel down the vertical axis. For instance, the majority of vehicles below the horizontal axis listed for more than $60 000 whereas the Expedition or Yukon listed in the mid $30K range at this time. The figure also reveals that brands sharing the same manufacturer are located close together. This means that consumers give these particular brands similar consideration in their purchase plans. Note that four of the GM vehicles are closely located in the first quadrant. Similarly, the two Cadillac Escalades, also GM vehicles, are grouped together, but given their premium pricing they are located with the higher priced vehicles. Other manufacturer groupings include the Ford Expedition and Ford Excursion, Toyota Land Cruiser and Toyota Sequoia, as well as the H1 and H2 Hummers. This particular joint space positioning map is problematic for this set of manufacturers in failing to make their own subset of brands sufficiently distinctive from each other and appealing to different sets of consumers in this particular product segment.

From Figure 3, we also note a heavier concentration of vectors in the top half of the graph near the Chevy and GMC products. We may conclude that for this vehicle segment, consideration to buy is greatest for the Chevy/GMC brands, followed by the Ford Expedition where we find the next largest concentration of consumer vectors. In this figure, note that a number of consumer vectors are superimposed on one another. This merely represents consumers with very similar consideration-to-buy ratings on all 16 brands. The lower part of Figure 3 shows much less segment consideration for the large luxury sport utility vehicles like the Range Rover and H1 Hummer.

The simple unfolding model results. For the simple unfolding model representation, we first reverse scaled the same input data ($5 - P_{ij}$) given Equation (2). Figure 4 displays the derived two-dimensional joint space spatial map for the simple unfolding or *ideal point* model. Recall that both brands and consumers are represented as points in this two-dimensional space. Here, brands are labeled and represented as diamonds for ease of identification and the 278 consumer are represented by points. The closer a consumer ideal point is to a brand, the higher the consideration to buy for that

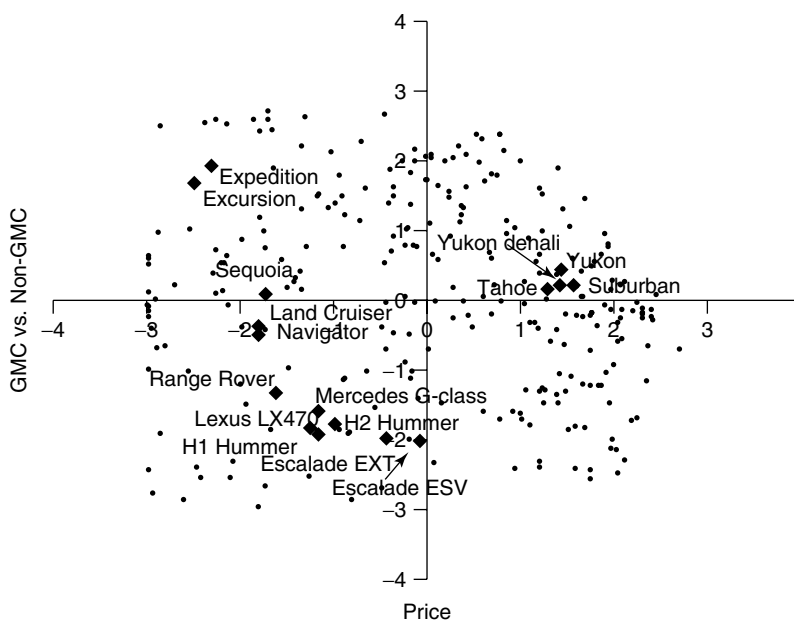


Figure 4 Simple unfolding model – SUV application.

particular brand. Figure 4 shows a configuration of brand locations similar to that in Figure 3, and we label these two dimensions accordingly, with Dimension I representing a GM versus non-GM factor, and Dimension II the price. As in Figure 3, the more expensive vehicles are located lower on the vertical axis while the GM vehicles are separated from the non-GM vehicles along the horizontal axis. As indicated above, brands sharing the same manufacturer are located near one another suggesting similar consideration to buy and similar positioning difficulties. The ideal point model shows a clustering of consumer points near the GM vehicles (Yukon, Yukon-Denali, Tahoe, and Suburban). There are also a good number of points in the fourth quadrant, the area between the Suburban and Tahoe and the premium-priced vehicles. Because of the Euclidean distance present in the interpretation of the unfolding model, we might interpret this as potential consumer demand for different varieties of large GM SUV's given the underlying dimensions which define the space.

Discussion. Both Figure 3 (vector model) and Figure 4 (unfolding model) depict a two dimensional joint space with 16 brands and 278 consumers. We find that the two figures are quite similar with respect to their respective brand configurations and interpretation of the underlying dimensions. Since MDS seeks to reveal the underlying structure in data, this result is not surprising given we use the same data matrix to estimate both models. Both representations reveal potential marketing positioning problems where each manufacturer's own brands appear to compete more against one another than against those brands of other manufacturers. More recent work in this area of MDS has attempted to simultaneously perform market segmentation and product positioning (*see* LATENT CLASS AND FINITE MIXTURE MODELS) given the same input data as illustrated with these two MDS models (DeSarbo, Manrai, and Manrai, 1994). In such latent structure MDS models, market segments replace individual consumers in such joint space representations.

Bibliography

Bennett, J.F. and Hays, W.L. (1960) Multidimensional unfolding: determining the dimensionality of ranked preference data *Psychometrika*, **25**, 27–43.

Borg, I. and Groenen, P.J.F. (2005) *Multidimensional Scaling*, 2nd edn, Springer, Mannheim.

Carroll, J.D. (1980) Models and methods for multidimensional analysis of preferential choice (or other dominance) data, in *Similarity and Choice* (E.D. Lantermann and H. Feger), Hans Huber Publishers, Bern, Stuttgart, Vienna, pp. 234–289.

Carroll, J.D. and Arabie, P. (1980) Multidimensional scaling. *Annual Review of Psychology*, **31**, 607–649.

Coombs, C.H. (1950) Psychological scaling without a unit of measurement. *Psychological Review*, **57**, 148–158.

DeSarbo, W.S., Grewal, R., and Scott, C.J. (2008) A clusterwise bilinear multidimensional scaling methodology for simultaneous segmentation and positioning analyses. *Journal of Marketing Research*, **45**, 280–292.

DeSarbo, W.S., Manrai, A., and Manrai, L. (1994) Latent class multidimensional scaling analysis, in *Advanced Methods of Marketing Research* (ed. R. Bagozzi), Blackwell Publishers, pp. 190–222.

DeSarbo, W.S. and Rao, V.R. (1986) A constrained unfolding model for product positioning. *Marketing Science*, **5**, 1–19.

Slater, P. (1960) The analysis of personal preferences. *The British Journal of Statistical Psychology*, **13**, 119–135.

Tucker, L.R. (1960) Intra-individual and inter-individual multidimensionality, in *Psychological Scaling: Theory and Applications* (eds. H. Gullikson and S. Messick), Holt, Rinehart, & Winston, New York.

discriminant analysis for marketing research applications

Fred M. Feinberg

Introduction

Among marketers' main tasks is *segmentation*: breaking consumers, products, and firms into meaningful groupings. As opposed to a continuum – for example, how much of a particular product a household consumes in a year – data often appear to the marketing analyst in discrete buckets, like "light," "medium," and "heavy" users. To aid in conceptual clarity and practical targeting, marketers want to understand and predict which consumers (or products, or firms, etc.) fit into which group. That is, we would like to use a set of variables at our disposal to help explain how, and perhaps why, these groupings come out the way they do, which in turn implies explaining patterns of *group membership*. Other examples abound. For instance, can we: easily distinguish new versus returning customers? explain which of our five stores various customers chose to shop at? understand how contracts get assigned to different countries, or regions in a distribution zone? find some way to entice "loyal" versus "switcher" customer groups? In each of these cases, we have a set of distinct objects – sometimes two (heavy vs light users), sometimes more (which brand was purchased?) – that constitute the *outcome* or the dependent variable. Our goal as marketing researchers is to understand how *predictors* at our disposal relate to that outcome. Just like in ordinary least squares (OLS) regression, these predictor variables can be interval-scaled (or binary, with certain caveats discussed later), but the dependent variable is what statisticians call "discrete" (or "nominal," "categorical," "polytomous," or "multinomial," depending on context; *see* MODELS FOR CATEGORICAL DEPENDENT VARIABLES).

And so we seek a tool that operates like ordinary regression in power, ease-of-use, and applicability, but whose goal is to *discriminate* among two or more known groups. As its name suggests, discriminant analysis (DA) is that tool. It builds on a foundation of ordinary MULTIPLE REGRESSION, and uses some ideas from other topics covered in this volume, like factor analysis (*see* EXPLORATORY FACTOR ANALYSIS), CLUSTER ANALYSIS, and MANOVA (*see* ANALYSIS OF VARIANCE AND COVARIANCE; REPEATED MEASURES ANOVA). It also stands in for formal discrete choice models like logistic regression (*see* LOGIT MODEL), which can perform many of the same tasks, but makes different assumptions about the nature of the data and (unobserved) errors.

A common convention is to distinguish DA, which discriminates between exactly two groups, and multiple discriminant analysis (MDA), which handles more. Because the former is just a special case of the latter, we call them both DA unless clarity is compromised. Because discriminating between two groups is easier, both conceptually and visually, most of our development will deal with that case, branching out to multiple groups later in the article, and building on the two-case setup, for which all terminology is identical.

Why Use Discriminant Analysis at All?

Marketing researchers typically wish to use DA to accomplish several primary tasks, each of which we examine in detail. Here is a description of each, followed by a concrete example.

1. *Discrimination*: How can we determine linear combinations of the predictor (independent) variables that help best discriminate among *known* groups (dependent variable)? ("Which variables in our database (past sales, demographics, etc.) best explain which of our customers *did* renew their service contracts in the last year?")
2. *Classification/Prediction*: Given a new set of data points (individuals, items, units) whose group memberships are *unknown*, how would they be classified into the preestablished groups? ("Can we use the identified variables to predict which customers *will* renew their contracts in the coming year?")
3. *Verification*: Are there *significant differences* across the various groups in the (predictor variable) profiles of the individuals found

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

in them? ("Does the customer data we pay to collect truly help in distinguishing those who renew from those who do not, or could it be chance alone?")

4. *Influence*: Which of the predictor variables best account for between-group differences? ("Can we *order* our customer profile variables by how much each helps in separating those who renew from those who do not? Do some fail to help discriminate at all?")

We see here several different uses for, and subanalyses within, DA. Generally speaking, we want to use known group memberships to estimate "discriminant functions"; use those functions to make predictions; figure out whether the whole model is useful; and, if it is, clarify which parts of it (i.e., predictors) help the most. In this way, DA is like multiple regression, which yields, analogously: a linear function of the "independent" variables; a "prediction" or "fit" for each data point; an overall measure of how well the model is working (F-test); and separate measures for the significance of each variable (t-tests). The only real differences are that DA works with a categorical dependent variable and, when there are more than two groups to distinguish, produces several discriminant functions, not just one. Keeping the correspondence between regression and DA in mind proves helpful in understanding how it operates and in interpreting typical output.

An important consideration in using DA is that group membership must be known, with certainty, for the cases we use to calibrate the discriminant function. It is the analyst's task to *reproduce* that pattern of membership as closely as possible using the available predictors. This stands in contrast to methods like cluster analysis, which seek to *determine* the best way to group items based on a set of input variables, with no "correct" group memberships to compare with, and methods like MANOVA, which seek to use known group memberships to predict other information about the members of those groups (MANOVA, in fact, is mathematically identical to MDA, and is in some sense DA in reverse). It also stands in contrast to the so-called "interdependence methods," like factor analysis, which seek to interrelate a group of variables, not to use them to predict a designated dependent variable.

An Introductory Example

To guide our discussion of DA, let us first consider a concrete marketing problem. Firms selling via the web often wish to distinguish between first-time and repeat customers. First-time customers need to be drawn to the site in order for the business to expand and to help spot market trends. Repeat customers are the lifeblood of firms for whom customer "stickiness" – their willingness not to buy from rival retailers who may offer better pricing – is critical. A target firm may also want to know which type of advertising or promotions are especially effective for each group or market segment. So, a key issue is how to best discriminate between these two groups, new and repeat customers.

In our example, let us assume that we have data from a representative day of operations for a multiproduct e-commerce firm, consisting of a number of variables describing customers purchasing at least one item from the website; each customer's total purchases for that day are recorded, so each appears only once. The dependent variable is whether this was their first purchase or not. We restrict our analysis to variables that would be known to the retailer – those recorded as part of the browsing and purchasing experience – although in many real-world applications other variables, involving survey, behavioral, or complex demographic data might also be available. For the purposes of graphical illustration, we will use just two predictor (independent) variables;

- X_1 = Total dollar amount of purchase (in hundreds; AMT)
- X_2 = Total time spent shopping (in tens of minutes; TIME)

Purists might quibble with the fact that we are treating both variables as interval-scaled, even though they are necessarily positive. Such are the problems in applied research. Rare is the application where all variables neatly conform to the multivariate normality assumption. Experience suggests that the results of DA are broadly

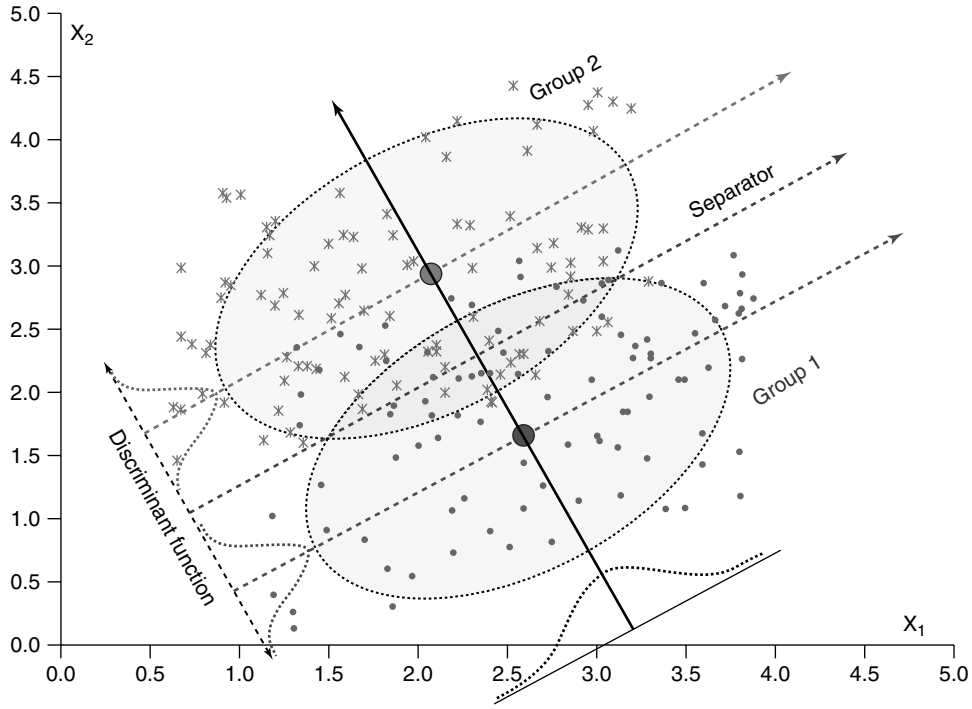


Figure 1 Geometric illustration of discriminant analysis for two groups, using two variables.

robust to this assumption, but we examine visual diagnostics and tests later that will help determine the extent of any problems its violation might cause.

Assumptions and Geometric Fundamentals

It is helpful to form a picture of what DA attempts to do, and then translate that picture into symbols. To do so, it is important to know what DA assumes about the structure of the data under consideration. We specify this first for the "two group" problem described above, and then generalize it for MDA. Precise, technical assumptions appear at the end of the article; what follows directly is a geometric and heuristic discussion.

DA's assumptions are familiar from both regression and, especially, ANOVA. If we have two groups to distinguish, we assume that the predictors (independent variables) have the same multivariate normal distribution *within each group*. Although this appears to rule out the use of binary predictors, it is usually fine to include a limited number of these as well, even if they do not strictly conform to the model's assumptions. In simple terms, each group's distribution differs only by its *mean;* once the mean is subtracted off, each group's data can be assumed to be drawn from the same normal distribution (bivariate for two predictors; multivariate for more). Graphically, for two variables and two groups (as in our marketing example), this looks like Figure 1.

In Figure 1, the data for both groups appear, superimposed with "80% probability" ellipses; that is, 20% of the points, on average, will fall outside (for more than two predictors, these will be ellipsoids in a higher-dimensional space; 80% was chosen for illustration purposes only). Note that these ellipses are presumed identical in shape, differing only in their mean, or *centroid*. An important point is that, if we assume the within-group data are multivariate normal, then *any* linear combination will be

Table 1 Group summary statistics.

Groups	*Variable*	*Mean*	*Std. Dev.*	*Group Size*
Group 1 "Repeat"	X1 (AMT)	2.64	0.759	100
	X2 (TIME)	1.89	0.719	100
Group 2 "New"	X1 (AMT)	1.90	0.749	101
	X2 (TIME)	2.79	0.699	101
Total	X1 (AMT)	2.27	0.840	201
	X2 (TIME)	2.34	0.838	201

(univariate) normal. The problem solved by DA is to find the one linear combination that minimizes the "spread" (standard deviation) of this common, within-group univariate normal distribution, relative to the between-group variation (where "variation" retains its usual statistical meaning as related to an appropriate sum-of-squares, corrected for interpredictor correlations, as discussed later). It is not difficult to discern that the linear discriminant function depicted in the graph accomplishes this; the line labeled "discriminant function" crosses the narrowest dimension of the cloud of data points. Put informally, sitting on a line in any other orientation would "see" more spread in the data: for example, the two superimposed normal distributions in the lower left "see" rather little within-group spread; the other superimposed normal, in the lower right, "sees" a great deal of spread. Note that any linear function of the two variables used in the space reduces the dimension of that space down to just one. In this sense, the discriminant function projects the space from multivariate to univariate, giving a single discriminant "score" to each point in each group, and minimizing the ratio of within-group to between-group variation.

Also depicted in Figure 1 is a "separator" line, which optimally classifies the points into the two groups. The classification in this case is not perfect: some points from each group would be classified as being in the other. This separator – which for MDA will be a higher-dimensional hyperplane orthogonal to the linear discriminant function – can also be used to *predict* group memberships for any new points requiring classification. As discussed later, this can be done using *classification functions* for each group.

Discriminant Analysis for New Versus Repeat Purchase Data

If we run DA on these data (depicted in Figure 1), we receive a sequence of results that tell us how well model assumptions are met, as well as various estimated quantities pertaining to discrimination. Let us examine selected output of a typical analysis. Before determining the discriminant functions, it is good practice to examine "summary statistics" for the data (Table 1).

Table 1 above tells us the groups are nearly equal in size (100 and 101, respectively); that group centroids appear different (2.64, 1.89 for Group 1; 1.90, 2.79 for Group 2); and that within-group standard deviations are roughly equal for variables X_1 = AMT (0.759, 0.749) and X_2 = TIME (0.719, 0.699). It is not entirely surprising that new customers spend, on average, less money ($190 vs. $264), but more time (27.9 vs. 18.9 minutes), on the site than repeat customers. Note that the standard deviations are greater for the entire group (0.840 for X_1, 0.838 for X_2) because these have not been "corrected" by the group centroids, so there is more overall variation. We next test to see whether the means for X_1 and X_2 are the same across groups (Table 2).

These tests, both "Wilks' Lambda," have significance levels below 0.0005, suggesting

Table 2 Tests of equality of group means.

	Wilks' Lambda	*F*	*df1*	*df2*	*Significance*
*X*1 (AMT)	0.803	48.876	1	199	0.000
*X*2 (TIME)	0.712	80.657	1	199	0.000

Table 3 Covariance matrices (before DA).

Groups		*X1 (AMT)*	*X2 (TIME)*
1	X1 (AMT)	0.577	0.230
	X2 (TIME)	0.230	0.516
2	X1 (AMT)	0.561	0.194
	X2 (TIME)	0.194	0.489
Total	X1 (AMT)	0.705	0.043
	X2 (TIME)	0.043	0.702

strong differences in the means across groups, and that discrimination should be powerful for these data. Next, we examine the *covariance structure* of the predictor data, both before and after DA is applied.

Table 3 looks at the data before any analysis. The top two sections compare the covariance for the two groups; they appear roughly equal. The lowest section, for the two groups pooled (taken together, or all 201 points), illustrates an interesting phenomenon: the covariance (and correlation) gets very small (0.043). The true correlation structure apparent in the ellipses of Figure 1 only shows up clearly when the groups are analyzed separately, that is, when they are mean-corrected.

Table 4 looks at the data after DA is applied, and a *common* covariance matrix is estimated for the two groups. Note that it appears to "lie between" the covariance matrices for the two groups (in the first table), and is close to an average of them. Finally, the correlation matrix tells us that, once we account for group differences, the correlation in AMT (X_1) and TIME (X_2) is about 0.397, which is sizable. This correlation is automatically accounted for in all "distance" measurements made in DA, as discussed more fully below.

Table 4 Pooled within groups matrices (after DA).

		X1 (AMT)	*X2 (TIME)*
Covariance	X1 (AMT)	0.569	0.212
	X2 (TIME)	0.212	0.502
Correlation	X1 (AMT)	1.000	0.397
	X2 (TIME)	0.397	1.000

Table 5 Test for covariance equality.

Box's M (F-test)	0.215
Significance	0.976

It is possible to rigorously test whether the two within-group covariance matrices indicate equality of covariance structure.

This test, Box's M, is an F-test, which the program assesses automatically as having a p-value of 0.976. Small values of p allow us to reject the null hypothesis of equal covariance matrices. Here, we cannot, meaning this assumption (covariance homogeneity) of DA is reasonable for these data (Table 5).

Next, we wish to determine which variables to include in the analysis (Table 6). The computer offers to do this for us using a *stepwise* procedure, explained more fully later in the article.

This part of the analysis indicates that both variables (X_1 and X_2) should be included in the discriminant function ($p < 0.0005$); apparently, of the two, X_2 should be "entered" first, as it does a better job discriminating than does X_1, if we could choose just one of them. That both variables should be used here is unsurprising; a quick glance at Figure 1 suggests as much. When there are more groups and more discriminant functions, this part of the analysis – *model selection* or *specification* – can be far more complex.

We now consider the discriminant function itself, and determine how well it helps distinguish the two groups.

We shall examine these tests and table entries in much more detail later. For the time being, the first (Table 7) tells us that the "canonical correlation" is 0.719, meaning that the DA analysis accounts for nearly 72% of the uncorrected variation in the total data. The second (Table 8) tells us that the significance, $p < 0.0005$, is strong; that is, the discriminant function explains a highly significant proportion of the variation in the data. Note that it is possible to have good (i.e., high) values of the correlation and bad (i.e., far from zero) values of p, and vice versa. The first measures the proportion of variation explained, which can be high even in small data sets; the second gives the likelihood this is by chance alone, and reflects how much data has been used in the analysis. In this way, they

Table 6 Determining variables to 'enter' the discriminant function.

			Wilks' Lambda: Exact F			
Step	*Entered*	*Statistic*	*Statistic*	*df1*	*df2*	*Significance*
1	$X2$ (TIME)	0.712	80.657	1	199.000	0.000
2	$X1$ (AMT)	0.483	105.892	2	198.000	0.000

are analogous to the R value and F-test in ordinary multiple regression. We next examine the discriminant function itself (Tables 9 and 10).

Table 9 shows the actual linear discriminant function: $f(X_1, X_2) = (-1.183)X_1 + (1.349)X_2 + (-0.578)$, where X_1 is the amount (in hundreds) and X_2 is the time (in weeks). We need not pay attention to the constant, which is there for technical reasons, much like in ordinary regression (note that, in Figure 1, we have used a different constant to make the graph easy to read, but the principle is the same). What the discriminant function says is that we get the "best" separation of the two groups when we multiply AMT by -1.138 and TIME by 1.349. Simply put, Group 1 ("Repeat") figures higher on AMT, and lower on TIME, than Group 2 ("New"). One might reply that this much was obvious from the outset, because the group means are ($264, 18.9 minutes) for repeat customers, and ($190, 27.9 minutes) for new customers (recall that our scaling are in hundreds of dollars and tens of minutes). But there are an infinitude of functions that would discriminate these numbers, including a vertical line (that is, $X_1 =$ AMT is constant) between 1.90 and 2.64, or a horizontal one ($X_2 =$ TIME is constant). The coefficients (-1.138, 1.349) not only distinguish these group means, but provide the *greatest possible discrimination (distance), relative to the remaining variation in the data*. In terms of marketing, if we needed to make a guess about whether a particular customer was a new customer or a repeat customer on the basis of their total purchase (AMT) and duration (TIME) alone, the single best index possible is the discriminant function above. Note that the function separates the group means well, with a value of -1.034 for Group 1 and 1.024 for Group 2. This distance (2.058) is large relative to the spread of the within-group data, which we know has variance of 0.569 for X_1 and 0.502 for X_2.

For completeness, we include two other sets of quantities closely related to the discriminant function above (Tables 11 and 12).

These are the *standardized* coefficients and the *structure matrix*. The first merely says what the DA function would be if the variables were standardized first (i.e., transformed so that their means are zero and standard deviations one, which is why there is no constant). The second lists the *correlation* between each variable (X_1 and X_2) and the DA function. Note that these are not the same as the standardized coefficients, because X_1 and X_2 are correlated with one another, within-group (with, as we saw earlier, an estimated correlation of 0.397). We will examine the standardized coefficients and structure matrix in more detail later in the article.

Finally, we would like to see how well the model works, in terms of classification, by which we mean the ability to predict which group a customer belongs to based on the amount and time alone. This is assessed in two ways, as shown in Table 13.

There are two matrices here, "original" and "cross-validated." The first shows that 85.0% of Group 1 and 83.2% of Group 2 cases were correctly classified by the DA analysis. The second recomputes the DA analysis with each of the 201 original cases left out, then reassigns that point based on the new DA coefficients (i.e., via 201 additional DA analyses, each using 200 cases). Although these results are identical here, this will not be the case in general. Overall, just over 84% (169/201) of the cases are correctly classified, which seem quite good relative to random guessing, based on the 100/201 and 101/201 "prior" probabilities, which reflect just over a 50% success rate.

We next present a general discussion of all the diagnostics, assumptions, metrics, and tests in both two-group DA and multigroup MDA.

Table 7 Eigenvalues.

Function	*Eigenvalue*	*% of Variance*	*Cumulative %*	*(Canonical) Correlation*
1	1.070a	100.0	100.0	0.719

Table 8 Wilks's Lambda.

Test	*Wilks' Lambda*	*Chi-square*	*df*	*Significance*
1	0.483	144.018	2	0.000

Table 9 Discriminant coefficients.

X1 (AMT)	–1.138
X2 (TIME)	1.349
(Constant)	–0.578

Table 10 Function at group centroids.

Group 1 (Repeat)	–1.034
Group 2 (New)	1.024

Discriminant Analysis in Detail

Two group and multiple group linear discriminant functions. When there are two groups, DA is closely related to multiple regression, if we allowed the dependent variable to be dichotomous. That is, the *linear discriminant function* is

$$L = b_0 + b_1X_1 + b_2X_2 + \cdots + b_mX_m \qquad (1)$$

where b_0 is a constant and $\{b_1, \ldots, b_m\}$ are coefficients for the predictors $\{X_1, \ldots, X_m\}$. Variables with the most significant (i.e., largest standardized) coefficients are most useful in group discrimination, via the linear function L.

For more than two groups, the picture is more complex. One could estimate a function like L above to discriminate the first group from all the others, or the last three groups from all the others, or use some other strategy. While this is possible, it is ad hoc, cumbersome, and would require a great deal of forethought from the researcher. Instead, it is far more sensible to use a function like L to provide the most *overall* discrimination amongst the entire set of groups; then a second function (orthogonal, or statistically uncorrelated, to the first) to account for as much *remaining* discrimination as possible; and then a third; and so on. This logic is analogous to the extraction of successive orthogonal factors in factor analysis, covered elsewhere in this volume (*see* EXPLORATORY FACTOR ANALYSIS). This procedure is computationally equivalent to *canonical correlation analysis*, and provides both the set of successively extracted linear discriminant functions and their *canonical roots* (which are, as in factor analysis, *eigenvalues*). These canonical roots or eigenvalues signify the relative discriminatory power of successive DA functions. Loosely speaking, a DA function with a root (eigenvalue) of 1.5 is 3 times better at overall discrimination that one with a root of 0.5. A useful, scale-free measure of DA function importance is supplied by dividing its root by the sum of all others; these, by construction, must add to 1 (i.e., 100% of the total discriminatory power of the model). The maximum number of discriminant functions is the lesser of {the number of groups minus one, the number of predictor variables}. Note that MDA and (principal components) factor analysis, while related, are *not* mathematically identical: the former maximizes the "distance" between values of the dependent (groupings) variable, whereas the latter maximizes the proportion of total variance accounted for by the factor solution.

Table 11 Standardized DA coefficients.

$X1$ (AMT)	–0.859
$X2$ (TIME)	0.956

Table 12 Structure matrix.

$X1$ (AMT)	–0.479
$X2$ (TIME)	0.616

Which variables help discriminate?

Interpreting coefficients. A key task in translating computer output to marketing understanding is interpreting the various coefficients provided. All programs provide, if asked, both unstandardized and standardized coefficients, b_i, for every predictor in each discriminant function. Standardized coefficients are generally preferred, as they neutralize scaling differences among the predictors (i.e., an analysis run with X_1 measured in feet versus inches would yield unstandardized coefficients 12 times larger for the first, but the same standardized coefficients). Just like in multiple regression, the larger its (standardized) coefficient, the greater the contribution that variable makes to discriminate among the groups. It is important to note that these coefficients speak *overall* discrimination, that is, among all the groups; they may not do a particularly good job at, say, distinguishing group 2 from group 5. To do that, the researcher can simply compare the *mean* of a particular function across the groups in question, or look at plots of one, two, or even three discriminant function means at the same time (more than this requires visualization of more than three dimensions).

Testing discrimination significance. It is useful to know whether a particular discriminant function in MDA is "significant" in helping to discriminate among the known groupings. *Wilks' lambda* assesses this significance, based on the eigenvalue for each function. Lambda takes values from 0 to 1, with *lower* values denoting greater discriminatory power. That is, values near zero suggest group means strongly differ (i.e., the function in question distinguishes among the set of groups), while values near one suggest the group means are about the same (for that function). A significant lambda allows one to reject the null hypothesis that the groups have the same mean discriminant function score, that is, the model does aid in discriminating. There are several other metrics besides Wilks' lambda to gauge how well DA discriminates between groups, among them being Rao's V and Mahalanobis' D-Square, each with an associated significance level. Any of these can be used as criteria for adding and removing variables in stepwise DA. In a similar vein, the *canonical correlation*, R*, measures how strongly the groups associate with the derived discriminant function. R* lies between 0 and 1; values near zero indicate a weak relationship, and values near one a strong one. It is also possible to compute an ANOVA table to test the overall adequacy of the DA model. The table yields an F-test to assess whether the model with a specific set of predictors discriminates better than a null model (i.e., one with no predictors, just a "grand mean", or constant).

Interpreting factor structure. Part of the problem in interpreting the (standardized) coefficients is that the predictor variables inevitably intercorrelate. It is possible, for example, that a particular variable appears to have very little influence, because its coefficient is small compared with that of another variable with which it is correlated; this *multicollinearity problem* complicates

Table 13 Classification results.

		Group	*Predicted Group Membership*		
			1	*2*	*Total*
Original	Count	1	85	15	100
		2	17	84	101
	%	1	85.0	15.0	100.0
		2	16.8	83.2	100.0
Cross-validated	Count	1	85	15	100
		2	17	84	101
	%	1	85.0	15.0	100.0
		2	16.8	83.2	100.0

interpretation in all regression-based methods. A distinct, and often more substantively useful, way to determine which variables "drive" a particular discriminant function is to simply compute the correlation between that variable and the function in question. This is called the *factor structure*, after factor analysis (where analogous quantities are called "*loadings*"). The key point here is that coefficients tell how much a particular predictor explains discrimination *over and above* other predictors in the model; by contrast, factor structure coefficients get at the full effect of each of the predictor variables, in a more "substantively meaningful" way. Therefore, DA coefficients represent *partial* correlations, while structure values represent *full* correlations. Both should be considered in interpreting any DA output. Note as well that the analysis program will provide "significance" levels for all coefficients and structure values; the researcher should strongly consider dropping anything nonsignificant from the final model, once an appropriate significance cutoff (α) is set.

Using DA for classification. Thus far, we have only discussed DA in terms of discrimination among *known* group memberships. Yet, marketing researchers will frequently wish to know how to *predict* which categories new points (e.g., customers, firms, products) fall into. Fortunately, it is a simple matter to use the output from DA to classify such new cases into the preestablished groups. Researchers must always bear in mind that any statistical procedure will typically fit better on the sample on which it was estimated/trained than it will on a new sample. That is, within-sample prediction can "overfit" idiosyncratic elements of training data, and go awry on a new sample with its own idiosyncracies. As such, nonparsimonious models (those including many predictors) can often appear to do a terrific job discriminating the training (calibration) sample, then do surprisingly poor on a new (prediction) sample; models with nearly as many predictors as data points can fit the training sample arbitrarily well, yet say little about prediction samples. Trying out one's estimated model on different samples is a form of "cross-validation," and is considered essential in assessing the validity of a particular model. There are several types of cross-validation available in major statistical programs. A popular option is *leave-one-out* classification, which involves rerunning the DA without a particular case and determining how it would be assigned, and performing this analysis for each case. In this way, it is rather like a *bootstrapping* procedure. It is particularly informative to the researcher, as it can help identify which cases seem to exert strong influence on the discriminant functions and their effectiveness.

Classification functions. Along with coefficients, standardized coefficients, and factor structure values, the DA program will also provide *classification functions*. These are conceptually and computationally distinct from the discriminant functions on which they are based. The purpose of the classification functions is, as one might guess, to classify new cases into the various preestablished groups, with one classification function per group. Deciding on the classification is simple: given a new observation, r, with coordinates $\{X_{1r}, \ldots, X_{mr}\}$ for its m predictors, we compute the classification score for each group and assign the observation to the group with the highest score. It is possible to use information about the *a priori* group membership probabilities as well, in line with so-called Bayesian reasoning. For example, if we know from vast experience that 63% of a site's visitors are male, we should somehow include this information in guessing at a new customer's gender. The value 0.63 is then the *prior probability*, and the value obtained from the model, conditional on this value, is called the *posterior probability*. Many computer programs allow the researcher to supply this information to be used in all calculations. As a final note, in some situations, a researcher may wish to use *K-nearest neighbor* (KNN) DA. KNN assigns a new case based on the groups to which its k-nearest neighbors belong, and can be a useful adjunct to the standard form of DA.

A word about "distance". Throughout, we have discussed minimizing and maximizing various "distances," without precisely defining what this means. Ordinary Euclidean distance, taught in high school geometry, is fine when our predictors are uncorrelated. When they are,

we must instead use the *Mahalanobis distance* instead. When all interpredictor correlations are zero – an extremely unlikely event for real data – then the Mahalanobis distance is the same as the Euclidean distance. When any of these intercorrelations is nonzero, the Mahalanobis distance will take it into account in calculating "distances" between points. When computing, for example, the distance between a particular point and the various group centroids, the Mahalanobis distance is always used. Note that Mahalanobis distance has an appealing scaling based on standard deviations from a group's centroid: for example, a point with (Mahalanobis) distance of two has a roughly 4.5% probability of belonging to that group, based on the standard z table.

Assessing the predictive accuracy of the model. There are many summary statistical measures to assess the performance of a particular DA model, among them being F-tests (Wilks' lambda), sums-of-squares, and log-likelihoods. However, these do not speak clearly to managers or retailers. Instead, a simple and evocative measure is the model's "hit rate": how often is it right? A version of this that provides more, but equally comprehensible, information, is the "classification matrix," which, for each group, tells how many times items in that group were correctly classified. The classification matrix goes by many other names (e.g., confusion, assignment, or prediction; matrix or table). Regardless of its title, it is formed by the rows being observed groups and the columns being the predicted ones. That is, it provides quantities for "How often was an item in Group X categorized as being in Group Y?" This way, a researcher can see if there are groups among which the DA model has particular difficulty classifying. A "good" classification matrix has most of its weight on the diagonal, representing correct predictions. The hit rate, therefore, is merely the proportion of cases appearing on the diagonal. Generally speaking, it is impossible to say what constitutes a "good" or "acceptable" hit rate, as this is so strongly situation-dependent. It may be easy to categorize customers as to their gender, but extremely difficult as to the state they live in.

Stepwise discriminant analysis. A popular way of determining which predictor variables should be included in the final discriminant function(s) is to use a *stepwise* method, which helps determine when to add or remove predictor variables. There are several forms of stepwise DA, and criteria on which they are based. In *forward stepwise* DA, one starts with no predictors in the model. At each "step," all variables are examined to determine which contributes most to group discrimination. That variable is then included, and the process starts over. In *backward stepwise* DA, one starts with all variables in the model and, at each step, the variable that contributes least to group prediction discrimination is removed. Each process stops when, respectively, adding or removing variables is not justified by an appropriate significance test (as below). If both methods yield the same "final" model, a researcher can feel reasonably confident in having searched the model space (i.e., all possible predictor variable subsets) well. If the methods do *not* yield the same final model, the researcher should carefully examine discrepancies, to determine whether predictor variable intercorrelations were responsible, and make appropriate adjustments, based on the project's goals and the *meaning* of the predictors.

Adding or removing predictors. Both forward and backward stepwise DA are guided by significance tests, specifically, the "F to enter" and "F to remove" values reported by the statistical program. These F values correspond to significance in group discrimination. In short, a variable that aids in group discrimination should be added, and one that does not should be dropped. Researchers must be cautious, however, because they are working with a *sample*, and stepwise procedures can often make variable selections based on chance features of the sample; that is, the training data in the sample may be overfitted. One way to guard against this is to set a higher alpha value for inclusion in the model; this will tend to result in models with fewer predictors. Many statisticians warn against using stepwise regression methods for this reason, but they can be a useful first pass for determining the best DA solution.

Some Additional Issues

Use of "dummy" variables. Strictly speaking, dummy variables, because they do not have a multivariate normal distribution, should not be used in DA. In practice, however, they are, and tend not to cause major problems if used judiciously. However, note that the influence of dummy variables cannot be judged by their individual (standardized) coefficients; rather, they must be assessed as a group. A simple way to do this is to run DA both with and without the set of dummies, and to compute the difference in squared (canonical) correlation, just as one compares R^2 values in two (nested) ordinary regressions.

DA versus logistic regression. Logistic regression was developed after DA. It is often used for the same ends, as it entails fewer assumptions and, consequently, fewer violations (specifically, we need not assume that predictor variables are interval-scaled, identically distributed, or are within-group multivariate normal). Furthermore, logistic regression is more robust in practice, in that group sizes need not be roughly equal. As in typical statistical applications, however, DA is more powerful – that is, the chances of not rejecting a false null hypothesis are less – when its (admittedly, many) assumptions are met. Logistic regression is nonetheless considered more "modern," and many researchers prefer its built-in flexibility.

Appendix: Assumptions for MDA

Categorical dependent variable. The dependent variable falls into "natural" discrete categories, not ones artificially created from underlying (perhaps latent or unknown) variables on another scale, like interval. In practice, the use of MDA is fairly robust to this assumption.

Interval-scaled predictors. All independent variances should be interval-scaled: continuous and without upper or lower bounds. In practice, this is often violated, and binary (e.g., dummies) or multipoint scales (e.g., Likert; *see* ITEMIZED RATING SCALES (LIKERT, SEMANTIC DIFFERENTIAL, AND STAPEL)) can be used with caution, much like they are in OLS-based multiple regression.

Homogeneous within-group multivariate normality. Within each group, the independent variables should have a similar *structure*: they can, of course, differ in means, but the within-group covariance matrix should be similar (the judged similarity will depend on the relative sizes of the groups; this can be assessed visually, via histograms of frequency distributions, or via formal tests, notably Box's M). DA can still be used if this is moderately violated, but extreme outliers should always be scrutinized carefully. Note that this multivariate distribution cannot be *degenerate*, which will occur if one independent variable is nearly a linear combination of others, causing "conditioning" problems.

Independence of observations and of errors. The values for any data "point" cannot offer information about any other. This means that, in consumer purchase data (for example), we cannot have some consumers appearing more than once, which could induce various correlations. All errors (distances) are independently and identically distributed (as a univariate normal, with zero mean an estimated standard deviation).

Model specification. The correct predictor variables have been entered into the model. Note that, if these are substantially intercorrelated, adding or removing variables can affect coefficient values and their significance levels.

Uncorrelated means and variance. A serious problem occurs with the accuracy of significance tests if predictor means are correlated with predictor standard deviations across groups. This can happen in practice if there are even a few extreme outliers; these can inflate the estimated within-group variances and, to a lesser extent, the pooled variance estimates. In such cases, outliers need to be examined and, if justified, removed.

Sample and group sizes. Various requirements dictate when a solution is possible, but these can be extreme in practice, and rules-of-thumb are better guides for the researcher. Strictly

speaking: each group must have at least two points (so that a centroid and variance can be computed); each independent variable has nonzero variation in each group; the maximum number of predictors is the number of data points, minus two; the maximum number of discriminant functions is the minimum of {the number of groups minus one, the number of variables in the analysis}. In practice: no group size should be more than 10 times greater than any other; there should be at least 5 times as many data points as predictor variables; each group should have more than 10 cases (but 20 is preferred, especially as the number of predictors grows). Researchers must bear in mind that these rules do *not* come from theory, but from vast experience with using the method on real and simulated data. In this sense, they are "best practices," and deviations are inevitable in applied research.

Acknowldgments

The author wishes to thank Elea Feit, Wagner Kamakura, and Carolyn Yoon for their feedback and constructive advice on this article.

Bibliography

Crask, M.R. and Perreault, W.D. (1977) Validation of discriminant analysis in marketing research. *Journal of Marketing Research*, **14**, 60–68.

Dillon, W.R., Goldstein, M., and Schiffman, L.G. (1978) Appropriateness of linear discriminant and multinomial classification analysis in marketing research. *Journal of Marketing Research*, **15**, 103–112.

Feinberg, F.M., Kinnear, T.C., and Taylor, J.R. (2008) *Modern Marketing Research: Concepts, Methods, and Cases*, Thomson Publishing, Mason.

Huberty, C.J. (1994) *Applied Discriminant Analysis*, John Wiley & Sons, Inc., New York.

Klecka, W.R. (1980) *Discriminant Analysis*, Sage Publications, Thousand Oaks.

Lachenbruch, P.A. (1975) *Discriminant Analysis*, Hafner Publishing, New York.

McLachlan, G.J. (2004) *Discriminant Analysis and Statistical Pattern Recognition*, John Wiley & Sons, Inc., New York.

Wedel, M. and Kamakura, W.A. (2000) *Market Segmentation: Conceptual and Methodological Foundations*, 2nd edn, Kluwer Publishing, Dordrecht.

multiple regression

Sangkil Moon

Introduction

Let us say that a marketing researcher wants to learn about the type of customers who want to buy an iPod touch (32GB for $399). For the analysis task, probably, the researcher will look at potential buyers' demographic information (e.g., gender, age, income), lifestyle information (e.g., How often do you listen to music while you are on the go? Do you usually buy new high-tech products earlier than your friends?), and preferences for the MP3 player product (e.g., How important is the video function on the MP3 player to you? Do you intend to buy the iPod touch within the next six months?) The researcher can collect such information from potential customers using a survey method. Then, everything is put together for a *multiple regression* analysis to describe what type of customers are likely to buy iPod touch (the dependent variable) using the above-mentioned information (multiple independent variables). Thus, the power of the multiple regression analysis technique lies in effectively describing or predicting a focal concept as the single dependent variable (e.g., purchase intention based on a 10-point scale) using other relevant information as independent variables.

Since this regression technique describes the relationships between the dependent variable and the chosen multiple independent variables, it goes beyond simple descriptive analyses such as obtaining averages or frequency tables (*see* FREQUENCY DISTRIBUTION) of variables of interest. From a marketing practitioner's perspective, it would be much more effective to develop actionable marketing strategies based on an understanding of the relationships identified from a multiple regression analysis than simply examining multiple individual variables separately. Certainly, car dealers need to learn how their potential customers would react (dependent variable) to car price decreases (independent variable) in an extreme economic recession. Therefore, the main purpose of this article is to provide a general and basic overview of multiple regression analysis as a practically useful statistical technique for solving real-world marketing research problems. Since multiple regression is a standard procedure in various statistical package software programs such as SAS (along with the Enterprise Guide interface, www.sas.com), SPSS (www.spss.com), and Minitab (www.minitab.com), the essay focuses on conceptual and intuitive aspects of this statistical methodology along with its applications in marketing research.

Basics of Multiple Regression

Multiple regression analysis is a statistical technique that can be used to analyze the relationship between a single dependent variable and multiple independent variables. The objective of multiple regression analysis is to use the independent variables that are known to predict the single dependent variable selected by the researcher (Hair *et al.*, 2006). A simplified case of this multiple regression analysis is simple regression analysis that has only one independent variable explaining the dependent variable. However, this simple regression analysis is hardly useful for practical marketing research because there are always multiple factors (e.g., demographics, lifestyle, product preferences) that explain the central concept (e.g., purchase intention). Even so, we learn about useful and realistic multiple regression based on the concept and structure of simple regression (or bivariate regression, Burns and Bush, 2006; McDaniel and Gates, 2008) that focuses on the relationship between the dependent variable and a single independent variable because conceptually multiple regression extends simple regression to multiple relationships between the dependent variable and each of the multiple independent variables.

Formally, we can express the multiple regression model equation as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon \quad (1)$$

where Y is the dependent variable (response variable or criterion variable), X's are k independent variables (or explanatory variables or predictor variables), β's are k parameters to be estimated, and ε is a disturbance (or error) term. In applying this multiple regression technique to

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

solve practical marketing problems, it is important to understand the properties of each term in this equation. First, both Y and X are obtained from data and, are therefore, *known* variables. For instance, in a survey data context, each variable, whether dependent or independent, can directly come from each appropriate research question in the survey questionnaire. The dependent variable, however, must be a continuous variable (e.g., length, price, sales, customer satisfaction on a 10-point scale) because the statistical model is a *linear* regression as opposed to binary or multinomial logit regression (*see* MODELS FOR CATEGORICAL DEPENDENT VARIABLES) that can take categorical variables as its dependent variable. For instance, if the problem is one of consumers' choosing their favorite brands in some product category, the logit regression model is used as opposed to the linear regression model because the dependent variable is categorical (e.g., brand alternatives). By contrast, if the product sales is used as the dependent variable, the linear regression is used as opposed to the logit regression because the dependent variable is continuous (e.g., sales in dollar amounts) (Feinberg, Kinnear, and Taylor, 2008). By contrast, we can use both continuous and categorical variables as independent variables in multiple regression but need to reorganize the regression equation in such a way that we identify each category in the categorical variable as a separate binary independent variable with one of the categories as the reference (or base) category. These categorical independent variables are called *dummy variables* and explained in more detail in the section Dummy Variables, later in this article. The linear regression model assumes that ε, the disturbance term, follows a normal distribution, which makes the regression model statistically powerful and convenient.

Finally, β's are unknown parameters and need to be estimated by minimizing the sum of the squared errors expressed in Equation 2,

$$\text{Sum of squared errors (SSE)} = \text{Min} \sum (Y - \hat{Y})^2 \quad (2)$$

$$\hat{Y} = b_0 + b_1 X_1 + b_2 X_2 + \cdots + b_n X_n \quad (3)$$

where b's are estimates for β's. This estimation method is called the *ordinary least squares* (*OLS*) method. Conceptually, the estimation rule amounts to finding the best linear line that minimizes the errors between the estimated regression line and a collection of the actually observed data points.

As one can logically expect, not all the independent variables initially selected by the researcher would influence the dependent variable *significantly*. Therefore, the research needs to "trim off" independent variables that do not have a legitimate influence on the dependent variable. Theoretically, the statistical significance level that is used for the cutoff level for the trimming-off job can range from 0 to 100% and set by the researcher based on his or her substantive goals. And yet, the most common cutoff significance levels are 1, 5, and 10%. The lower the statistical significance level, the stricter is the test. In other words, a lower significance level results in fewer significant independent variables. To test such a *statistical significance* of each independent variable, we compute the p value for the independent variable, which also ranges from 0 to 100%. The p value measures the confidence that the analyst can have in the estimated value for the intercept or each independent variable in the regression model. The lower the p value for the independent variable is, the more significant is its estimated value. Then, the p value is compared with the predetermined cutoff significance level. If the p value for the particular independent variable is smaller than the predetermined significance level, the variable is determined to be statistically significant. (These p values are automatically reported along with parameter estimates on a statistical program.) The procedure of selecting a proper regression model with only significant independent variables is called STEPWISE REGRESSION. There are three approaches for conducting the stepwise regression – (i) forward selection, (ii) backward elimination, and (iii) hybrid approach (combining both forward selection and backward elimination). *Statistical significance* is required for a relationship to have validity, but statistical significance without theoretical support (*practical significance*) does not support validity. Basically, practical significance means that the result makes sense conceptually and theoretically but cannot be

tested statistically in a straightforward manner unlike statistical significance. To summarize, before we apply the results of any regression analysis to a real-world case, we need to confirm both types of significance.

Interpreting Regression Parameter Estimates

b_i is an estimate for parameter β_i. More specifically, b_i is our best guess for the *unknown* true parameter value β_i. The estimate measures both the *direction* and *size* of the impact of X_i on Y while holding all the other predictors constant. For example, let us say Y = total brand sales (in dollars) and X_1 = brand price (in dollars) along with other independent variables such as brand promotion frequency and brand consumer ratings being used. We obtain $b_1 = -500$ from our regression. What does that mean? First, the negative sign indicates the brand price increase has a negative influence on the sales of the brand (impact direction). Second, as the brand price is increased by \$1, the total brand sales decreases by \$500 (impact size). This can be very useful information for marketers because they can tell how much sales they will lose when they increase the price by a certain amount. There is, however, a caveat in using this kind of interpretation; the regression result is valid only within the observed ranges of the dependent and independent variables based on the linear relationship taken by the model. Specifically, beyond the ranges used in the data, the estimated impact direction and size can be misleading or false. Moreover, the actual relationship between the dependent variable and the independent variable may not be linear, which is not captured in linear multiple regression in a straightforward manner. On the other hand, the parameter estimate for the intercept (b_0) indicates the value of the dependent variable when all the independent variables take the value 0.

It should also be stressed that the estimates are scale dependent, which means that the scales used by the independent variable and the dependent variable can change the size of the estimate, while the actual relationship remains the same. For instance, if we change the scale of X_1 (brand price) from dollars to cents, b_1 becomes -5 from -500, while we maintain the scale of the dependent variable (unit dollar). This change means that as the brand price increases by 1 cent, the total brand sales decreases by \$5. In brief, even though the estimate value changed significantly, its real meaning and interpretation should remain the same since the change is only a reflection of a changed scale. Lastly, we often see that the researcher draws a causal relationship based on his or her regression analysis but it is a general understanding that regression alone does not prove any causal relationships.

The variations in variable scales and different variabilities across variables make a direct comparison of multiple independent variables problematic and challenging. Standardization converts each variable to a common scale and variability, the most common being a mean of zero and standard deviation of one. With this approach, we make sure that all variables are comparable on the same scale. The advantage is that it eliminates the problem of dealing with different units of measurement and thus reflects the relative impact on the dependent variable of a change in one standard deviation in each variable. Although this variable standardization allows the researcher to compare the independent variables on a level playing field, it can cause more problems in interpretation and is also limited in its practical uses. For instance, what is the meaning of one standard deviation increase in household income (X_1) as an independent variable that influences the customer's intention to purchase an iPod touch (Y) in comparison with one standard deviation increase in the product price (X_2)? First, the two independent variables are not quite comparable as marketing tools. Second, the meaning of one standard deviation is less clear than the meaning of one dollar as the unit of both variables.

Dummy Variables

It was mentioned above that, in multiple regression, whereas the dependent variable should be continuous, the independent variables can be continuous or categorical. In treating categorical independent variables, the researcher needs to code binary dummy variables differently from continuous variables in his or her regression analysis. In the regression data matrix, each row takes one observation of the analysis unit (e.g.,

survey respondent, client, firm), whereas each column takes one conceptual construct that can be an independent or dependent variable (e.g., each research question in a survey, each demographic variable). Continuous independent and dependent variables can take their own value, whereas each categorical variable must be set up differently into two or more dummy variables. If the categorical variable has only two categories (e.g., gender with female and male, yes–no question), one of the categories (e.g., male) can be coded as 0 arbitrarily and the other (e.g., female) can be 1, again arbitrarily. Then, the positive sign of its parameter estimate indicates that category 1 (e.g., male) has a positive relationship with the dependent variable against category 0 (e.g., female) as the reference or base category. In the same way, its negative sign indicates category 1 has a negative relationship with the dependent variable. Thus, the dummy variable can only be interpreted in relation to its reference category. When the categorical variable has more than two categories (e.g., three agents in online travel businesses), the coding scheme becomes more complicated by taking c–1 variable columns where c is the number of categories in the original variable. The number of dummy variables is one smaller than c because the reference category does not appear in the data matrix. More details on the dummy variable can be found in Churchill and Iacobucci (2005).

How to Choose Relevant Independent Variables?

The researcher wants to evaluate how good the multiple regression model specified and estimated is. Model specification indicates what independent variables are used to explain or predict the dependent variable of interest. There are various aspects of the model the researcher needs to evaluate to determine how good the model is.

One of such diagnostic measures is the coefficient of determination (R^2). R^2 is used to measure the proportion of the variance of the dependent variable explained by the set of independent variables used and ranges from 0 to 1. The higher the coefficient, the better the model is in terms of data fitting with 1 indicating the estimated regression line perfectly fits the data. There is no clear cutoff level in R^2 to accept or reject the estimated regression model even though novices quite often tend to use this value as the all-purpose diagnostic measure in evaluating the model under investigation. This can be misleading because the R^2 measure also depends on the sample size in association with the number of parameters being estimated. For example, a multiple regression with five predictors estimated on six data points will produce the R^2 of value 1 regardless of data or predictors. Importantly, however, if the measure is too low, it means the model, particularly the independent variables used, does not explain the dependent variable properly, which should caution the researcher to change the regression model by finding other independent variables that can explain the dependent variable better. It should be noted that the addition of more independent variables will always increase the R^2 value, even if nonsignificant ones are added. By contrast, with the adjusted coefficient of determination (adjusted R^2), adding nonsignificant independent variables just to increase R^2 is discounted in a systematic manner. In other words, when the researcher adds another independent variable, the adjusted R^2 will balance a positive factor (added explained variance by the added variable) with a negative factor (penalty for having another independent variable). Therefore, when we add another independent variable to our regression model, adjusted R^2 can go either way unlike R^2. For this reason, adjusted R^2 can be effectively used to determine whether the researcher needs to add a particular independent variable (Hair *et al.*, 2006).

Any estimation method for multiple regression including the above-mentioned OLS method requires that the chosen model is not misspecified. Misspecification usually concerns the notion that the maintained assumptions for the unobserved error variable ε are violated or that the functional form, which is linear, is inappropriate (Franses and Paap, 2005). The most problematic issue in independent variable selection is specification error, which concerns the inclusion of irrelevant variables in or the omission of relevant variables from the set of independent variables. We can find out whether a model is misspecified only after we conduct diagnostics testing for a variety

of theoretically possible models including the coefficient of determination test. As a rule of thumb, when it is unclear, the researcher would consider including potentially irrelevant variables rather than possibly omitting relevant variables. From a broad perspective, finding the properly specified multiple regression requires a lot of trial and error with both statistical skills and conceptual reasoning. Besides, this misspecification error concern related to the assumptions for the unobserved error variable in linear regression makes the analyst transform some limited continuous dependent variables such as the sales variable that can take only a positive value. In this example, since the limited range of the dependent variable can skew the distribution of the unobserved error variable in an undesirable fashion, the analyst would need to take the log value of the original sales variable as a transformed dependent variable. To specifically transform the dependent variable a careful inspection of the residuals $(Y-\hat{Y})$ defined by Equations 1–3 is required.

In determining the set of independent variables, the researcher also would need to check MULTICOLLINEARITY. Multicollinearity refers to the correlation among independent variables. Generally, the correlation indicates the relationship between two variables, whether positive or negative, which is commonly measured by Pearson's product-moment coefficient of correlation (*see* PRODUCT MOMENT CORRELATION) (Glass and Hopkins, 1984). For instance, in explaining consumers' general movie ratings (dependent variable), their evaluations of the subcategories of acting and directing (two independent variables) can be highly correlated. If the correlation is too high, we need to drop the inferior of the two variables from the independent variable set because the two variables play about the same role in explaining the dependent variable. Formally speaking, the impact of multicollinearity is to reduce any single independent variable's predictive power by the extent to which it is associated with the other independent variables (Hair *et al.*, 2006). To maximize the prediction from a given number of independent variables, the researcher should look for independent variables that have low multicollinearity with the other independent variables but also have high correlations with the dependent variable. To formally assess multicollinearity, we can use the tolerance measure expressing the amount of variability of the selected independent variable not explained by the other independent variables; this ranges from 0 to 100%. A high tolerance value means a small degree of multicollinearity, which is good for the model. A common but rough cutoff threshold is a tolerance value of 10%.

The most appropriate empirical validation approach is to test the regression model on a new sample drawn from the general population (validation sample or holdout sample) against the sample used for estimation (estimation sample or calibration sample). When we obtain similar results from both samples, we can be more confident that the regression estimates are valid. For the task, the researcher needs to set aside a proportion of the original sample as the validation sample that will not be used for the estimation purpose. One downside to this validation approach is it makes the estimation sample size smaller, which reduces the statistical power of the estimation approach in various ways. If the researcher has a large sample, this would not be a problem. A larger sample, however, increases the research budget.

Applications of Multiple Regression in Marketing Research

Multiple regression can be applied to cases where the researcher is interested in the relationships between a central concept (e.g., customers' purchase intention) and other concepts or measures (e.g., age, product rating, previous purchase experience) that can potentially explain the central concept. The analysis technique has an extremely wide applicability and has been applied in most disciplines including marketing research. Most marketing research books introduce the technique in varying degrees of detail with practical application examples. Several such examples from some popular marketing research textbooks are briefly introduced here.

First, Parasuraman, Grewal, and Krishnan (2004) used an example of a school librarian who is eager to increase the number of students borrowing books from the library. They indicated that she can use multiple regression to

provide persuasive evidence that demonstrates how increased borrowing of books can benefit students by explaining students' GPAs (dependent variable) with number of books borrowed (independent variable). Second, Cooper and Schindler (2006) introduced an example of applied multiple regression to evaluate the key drivers of customer usage for mail (dependent variable). In the analysis, cost, speed, security, and reliability of mail turned out to be significant predictors of mail usage. Third, Hair *et al.* (2008) provided a regression example, where the researcher applies multiple regression to examine the relationship between perceptions of the food in the restaurant (independent variables) and overall customer satisfaction (dependent variables). Food perceptions included customer ratings of the restaurant on fresh food, food taste, and food temperature. Lastly, Malhotra (2009) discussed an example of Wendy's customers' patronage likelihood (dependent variable) in multiple regression. Customer ratings of Wendy's on menu selection, quality, service, value, and location convenience were used as independent variables to explain its customers' patronage likelihood.

Warnings in Using Multiple Regression

As shown everywhere, including real-world marketing research, multiple regression is one of the most popular and powerful tools from among numerous common statistical analysis tools such as correlation and analysis of variance (ANOVA). As in any analysis tool, however, it has its limitations, which should be considered before the researcher decides to apply the technique to his or her own problem. Several common caveats are described here.

First, multiple regression basically captures only a *linear* relationship between each independent variable and the dependent variable as a type of *linear* regression. Even though the linear relationship is a fundamental assumption, it is possible to capture some limited types of curvilinear relationships by transforming or extending original variables. For instance, the theory we want to test says that there is an inverted U relationship in advertising spending (independent variable X_1) on consumers' brand image (dependent variable). In this case, the inverted U relationship means that advertising spending enhances consumers' brand image only to a certain level of advertising when the advertisement can raise consumers' awareness of and familiarity with the advertised brand. Beyond the saturation level, however, the same advertisement can do more harm than good to its brand image because too much advertising can elicit negative emotional reactions (e.g., boredom, fatigue) in consumers. To test the theory, we can add the quadratic term of the original linear variable (independent variable X_1^2) to the regression equation while still keeping the other variables. The combination of both parameter estimates will determine whether there is an inverted U relationship between advertising spending and consumers' brand image. More general curvilinear relationships can be captured by extending the original independent variable to a higher degree of polynomial terms (e.g., X_1^3, X_1^4).

Second, the ranges of data on the dependent and independent variables influence how widely the regression results can be applied (Parasuraman, Grewal, and Krishnan, 2004). In principle, the regression results apply only within the range of the data used for estimation. In the above example on the influence of advertising spending on brand image, let us say the collected data covered only the range between \$1 million and \$10 million. Then, we cannot directly apply the parameter estimate result to the case of \$20 million, which is beyond the range of the data used.

Third, the relationship that is identified as statistically significant in regression does not prove a cause-and-effect association (*see* CAUSAL RESEARCH). The statistically significant relationship in regression simply indicates an association between the independent and dependent variables without any direction. In a simple linear regression, if we change the roles of the independent and dependent variables, the regression will generate similar results because the method itself is blind to the direction of the impact. Generally, the cause-and-effect association cannot be proved by statistical methods alone. The researcher should use his or her own reasoning and intuitions based on solid theories and practical examples to prove a causal relationship.

Bibliography

Burns, A.C. and Bush, R.F. (2006) *Marketing Research*, 5th edn, Pearson Prentice Hall.

Churchill, G.A. and Iacobucci, D. (2005) *Marketing Research: Methodological Foundations*, Thomson South-Western.

Cooper, D.R. and Schindler, P.S. (2006) *Marketing Research*, McGraw-Hill Irwin.

Feinberg, F.M., Kinnear, T.C., and Taylor, J.R. (2008) *Modern Marketing Research: Concepts, Methods, and Cases*, Atomic Dog.

Franses, P.H. and Paap, R. (2005) *Quantitative Models in Marketing Research*, Cambridge University Press.

Glass, G.V. and Hopkins, K.D. (1984) *Statistical Methods in Education and Psychology*, Allyn and Bacon.

Griffiths, W.E., Hill, R.C., and Judge, G.G. (1993) *Learning and Practicing Econometrics*, John Wiley & Sons, Inc.

Hair, J.F., Black, W.C., Babin, B.J. *et al*. (2006) *Multivariate Data Analysis*, 6th edn, Pearson Prentice Hall.

Hair, J.F., Wolfinbarger, M., Ortinau, D.J., and Bush, R.P. (2008) *Essentials of Marketing Research*, McGraw-Hill Irwin.

Malhotra, N.K. (2009) *Basic Marketing Research: A Decision-Making Approach*, 3rd edn, Prentice Hall.

McDaniel, C. and Gates, R. (2008) *Marketing Research Essentials*, John Wiley & Sons, Inc.

Parasuraman, A., Grewal, D., and Krishnan, R. (2004) *Marketing Research*, Houghton Mifflin Company.

William, G.H. (2003) *Econometric Analysis*, 5th edn, Prentice Hall.

structural equation modeling

Hans Baumgartner

Introduction

Structural equation modeling (SEM) with latent variables, also known as *latent variable path analysis*, is a technique for investigating relationships between latent (unobserved) variables or constructs that are measured by multiple manifest (observed) variables or indicators. Special cases are CONFIRMATORY FACTOR ANALYSIS (*see* EXPLORATORY FACTOR ANALYSIS), in which no directional relationships between latent variables (factors) are specified, and manifest variable path analysis, in which directional relationships between observed variables are modeled. As an illustration, consider the model shown in Figure 1. This model is an attempt to explain consumers' usage of coupons for grocery shopping via a number of antecedent constructs. According to the model, actual coupon usage is a function of consumers' intentions to use coupons, which in turn depend on their attitudes toward using coupons. Attitudes are influenced by three types of beliefs about the consequences of using coupons: rewards (positive consequences associated with coupon usage such as saving money on the grocery bill); inconveniences (one type of negative consequence associated with coupon usage such as the time and effort required to use coupons); and encumbrances (another type of negative consequence associated with coupon usage such as the need to buy nonpreferred brands in order to take advantage of coupon offers).

This simple model illustrates two of the major advantages of SEM. First, SEM makes it possible to study complex patterns of relationships among the constructs in one's model in an integrative fashion. If regression analysis (*see* MULTIPLE REGRESSION) were used to analyze the model in Figure 1, one would have to estimate three different regressions, and it would be quite cumbersome to show that attitudes and intentions mediate the effects of beliefs on coupon usage (i.e., that there are no direct effects of beliefs on intentions and coupon usage). In contrast, all relationships among model constructs can be investigated simultaneously using SEM, and it is straightforward to examine different forms of mediation in complex multistage models. Furthermore, tests of overall model fit are available, which indicate how well the specified model represents the data. Second, constructs such as beliefs, attitudes, and intentions are not directly observable. Using single-item measures to assess the constructs of interest (e.g., asking people whether their attitude toward using coupons is favorable or unfavorable) does not do justice to even relatively simple constructs such as coupon attitudes (the situation becomes more problematic when a construct is more abstract and when it has many facets), and combining multiple measures into averages is ad hoc and makes assumptions that cannot be verified. Furthermore, observed variables always contain error, both random and systematic, and this measurement error has to be taken into account in the analysis because it can have serious distorting influences on investigations of relationships between the constructs in the model. Although SEM is not a panacea for these problems, it encourages researchers to think more explicitly about measurement error, enables assessment of the quality of construct measurement (*see* VALIDITY AND RELIABILITY), and facilitates the study of structural relationships between constructs in the presence of measurement error.

In this article, we provide an overview of SEM. The presentation is organized around the steps involved in using SEM for theory testing: model specification; preliminary data analysis; model estimation; overall model evaluation; model modification; and model interpretation (Bagozzi and Baumgartner, 1994; Baumgartner and Homburg, 1996). The article concludes with a brief discussion of more advanced uses of SEM and recent developments.

Model Specification

Graphical model specification. Models can be specified graphically or algebraically, but we focus on graphical specification, because many researchers find it more revealing. Figure 2 shows the graphical specification of a particular version of the model in Figure 1, assuming that two indicators each are available to measure

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

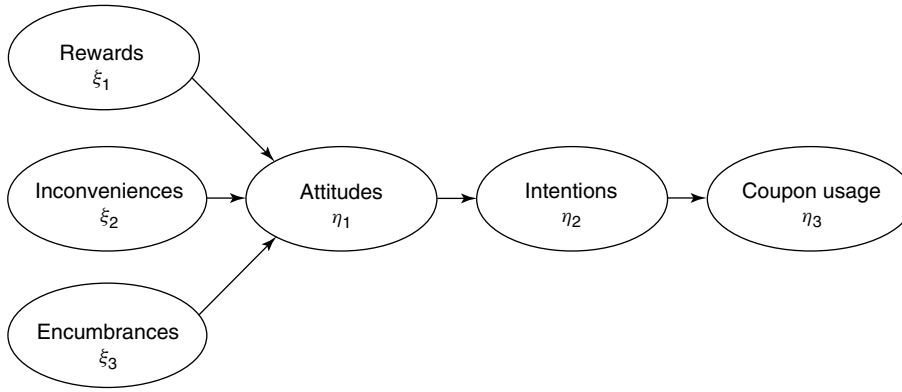


Figure 1 Theoretical model of the antecedents of coupon usage.

rewards, inconveniences, and intentions, three indicators to assess encumbrances, and four indicators to capture attitudes. Coupon usage, although shown as a latent variable, is actually treated as an observed variable. Latent variables that are of theoretical interest are shown as ellipses (or sometimes as circles). Attitudes, intentions, and coupon usage are referred to as *endogenous latent variables* (denoted by the Greek symbol η) because the model is designed to explain the variation in these variables. Rewards, inconveniences, and encumbrances are called *exogenous latent variables* (denoted by ξ) because they are not explained within the context of the model. The exogenous latent variables are usually allowed to covary freely, and these covariances are shown as double-headed arrows (denoted by φ_{ij}). The relationships of primary theoretical interest are the effects of the exogenous variables on the endogenous variables (denoted by γ) and the effects of the endogenous variables on each other (denoted by β). For example, γ_{11} refers to the effect of rewards on attitudes, and β_{21} to the effect of attitudes on intentions. Associated with each endogenous latent variable is an error term (error in equation or equation disturbance, denoted by ζ), because it cannot be expected that the antecedents of the latent variable can explain it completely. Arrows that emanate from and point to the same variable are variances, and the variances of the ξ's and ζ's are called φ_{ii} and ψ_{ii}, respectively (the ζ's are assumed to be uncorrelated in the present case, although that is not always necessary). The relationships among the exogenous and endogenous latent variables shown as ellipses constitute the so-called *latent variable model* (sometimes called the *structural model*), which represents the theoretical model studied in the research.

To examine the theoretical model empirically, each latent variable of interest has to have at least one observed measure. The indicators of the exogenous (endogenous) variables are called x (y), and by convention they are enclosed in rectangles (or squares). The coefficients relating the observed variables to their underlying latent variables (so-called factor loadings) are denoted by λ^x and λ^y. The error terms (errors in variables, unique factors) associated with the x's and y's are called δ and ε, respectively, and their variances are denoted by θ^δ and θ^ε. In the present case, the error terms were specified to be uncorrelated, but this assumption can be relaxed. The model linking the exogenous and endogenous latent variables to their observed measures is called the *measurement model*.

The model in Figure 2 is a complete specification of the structural equation model of interest, assuming that all relationships between variables are linear. The model contains five types of parameters: the factor loadings (λ^x and λ^y); the measurement error (unique factor) variances (θ^δ and θ^ε); the variances and covariances of the exogenous latent variables (φ); the latent variable model coefficients (γ and β); and the (co)variances of the errors in equations (ψ).

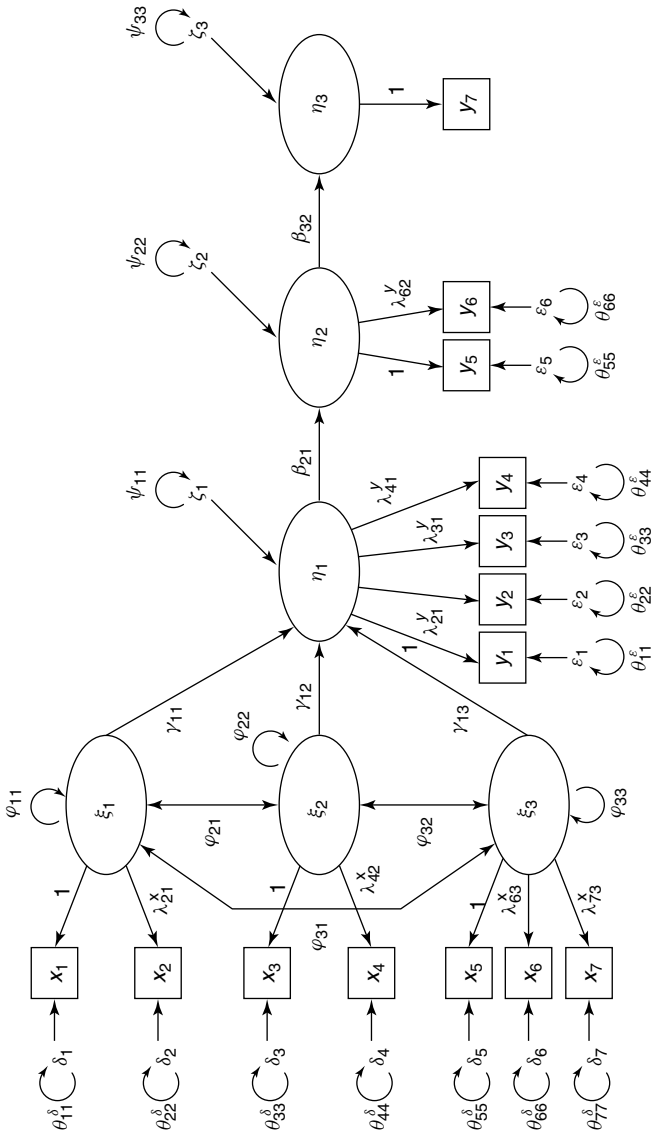


Figure 2 A structural equation model of coupon usage.

Model identification. For the model to be meaningful, it has to be identified. This means that all parameters in the model are uniquely determined so that the conclusions derived from the analysis are not arbitrary. As a first step, the scale of the latent variables has to be fixed (since the scale in which they are measured cannot be observed directly). One way to do this is to set one loading per latent variable to 1, as shown in Figure 2. In addition, the coefficients relating the error terms (both measurement errors and errors in equations) to observed or latent variables are set to 1 (not shown explicitly in Figure 2 because it is always done by convention). Furthermore, if there is only a single indicator, as in the case of coupon usage, measurement error is usually ignored and the error (unique) variance is set to 0 (another possibility is to assume a certain amount of error variance, preferably based on prior research).

A necessary requirement for identification is that the number of parameters to be estimated should not be greater than the number of unique variances and covariances among the observed variables. For the model in Figure 2, 35 parameters have to be estimated, and there are 105 unique elements in the observed covariance matrix. Therefore, the model has 70 degrees of freedom and the necessary condition for identification is satisfied (since the number of degrees of freedom is nonnegative).

Unfortunately, this simple rule does not guarantee identification. For certain types of models, it is easy to check identification (see Bollen (1989), for an extended discussion), but in general identification is a nontrivial issue. For general structural equation models consisting of both measurement and latent variable models the so-called two-step rule (which is sufficient but not necessary) is often applied. In the first step, the researcher tries to ascertain that a measurement model in which no restrictions on the latent variable model are imposed and all covariances between (exogenous, endogenous) latent variables are freely estimated is identified. If there are at least two indicators per construct, each indicator loads on one and only one construct, the measurement error terms are uncorrelated, and the constructs are allowed to freely correlate, the measurement model is identified. Even if a construct is measured by a single indicator, as long as the error variance is fixed to a certain value (usually 0, in which case the latent variable is really an observed variable), the model is identified. Using these guidelines, the measurement model corresponding to Figure 2 is identified (with 63 degrees of freedom).

Once the model in the first step is shown to be identified, the exogenous and endogenous latent variables can be treated as observed and the task in the second step is to show that the latent variable model is identified. If the model is recursive (i.e., the matrix containing the β coefficients is subdiagonal, which means there are no reciprocal paths or feedback loops so that latent variables later in the sequence do not influence latent variables earlier in the sequence, and the errors in equations are uncorrelated), the latent variable model is identified. It can be checked that the model in Figure 2 is recursive and therefore identified. If the model is nonrecursive, other rules, such as the rank rule sometimes used to identify systems of simultaneous equations, can be employed.

At times, it is too difficult to show identification explicitly and in that case one may have to rely on the computer program used to flag a nonidentified model. A useful strategy is to start with a simpler model that is known to be identified and to then complicate it by introducing, one by one, the additional parameters to arrive at the desired specification, provided the modification indices associated with the added parameters are nonzero (see the discussion of model modification later in the chapter).

Measurement model specification issues. In Figure 2, it is assumed that the observed variables are functions (or manifestations) of the latent variables. This is called a *reflective measurement model specification*. If an indicator is specified as reflective (i.e., as an *effect indicator*), one assumes that the different indicators of a construct are interchangeable, correlated with each other, and contaminated by measurement error. It is not always meaningful to regard indicators as manifestations of an underlying latent variable. Sometimes, it is more appropriate to conceptualize a given set of indicators as defining characteristics of a construct, in which case the latent variable depends on its

indicators. This is called a *formative measurement model specification*. Formative indicators (also called *cause indicators*) are not interchangeable, need not be correlated, and are usually assumed to contain no measurement error. Since statistical means of trying to distinguish between reflective and formative measurement models are of limited usefulness, the decision whether indicators should be modeled as reflective or formative has to be made based on a conceptual analysis of the items in question. The issue is important because research shows that incorrect specification of the measurement model can lead to significant bias in model parameters (see Diamantopoulos, Riefler, and Roth (2008), for a recent review and additional references). Although indicators should not be specified as reflective when they are not (e.g., satisfaction with different aspects of a job as measures of overall job satisfaction), there are many unresolved issues in formative measurement models, which makes their routine use problematic.

Another issue that has to be considered carefully is how many indicators of each latent variable should be included in the model and how these indicators should be related to the latent variable. In general, it is better to have more rather than fewer indicators per construct, but there are practical constraints on how many items can be put in a questionnaire. Furthermore, including too many individual indicators in a model might make the model too complex, and it is also difficult to get acceptable overall model fits when there are too many indicators per construct. An alternative is to use item parcels, where subsets of individual items are averaged and then used as (multiple) indicators in the model. This is acceptable if the items that are combined are known to be unidimensional (i.e., form a homogenous set), or if items measuring the same facet of a multidimensional construct are combined. If the factor structure of a set of measures is poorly understood, parceling is not a good idea (see Bandalos and Finney (2001), for additional discussion of item parceling).

Latent variable model specification issues. It is generally much more informative to propose alternative latent variable models that, based on prior research or new theorizing, could be plausible representations of the relationships among the constructs of interest than to present a single structural model that supposedly best represents the data. In other words, it is desirable that researchers adopt a model comparison approach to SEM (see MacKenzie, Lutz and Belch (1986), for a good example). Furthermore, researchers have to keep in mind that different specifications that can have very different substantive implications may fit the data equally or nearly equally well (MacCallum *et al.*, 1993).

Sample size. Researchers should make sure prior to data collection that the sample size will be adequate to avoid estimation problems and to get reliable fit statistics, parameter estimates, and estimates of standard errors. Although many kinds of factors can be expected to influence the required sample size, two kinds of heuristics for sample size determination have been proposed in the literature. Absolute guidelines are based on the notion that the sample size should be greater than a certain minimum number (e.g., there should be at least 200 observations). Relative guidelines specify that the number of observations per parameter estimated should exceed a certain minimum (e.g., ratios of 5 : 1 to as high as 20 : 1 have been mentioned). If sufficient prior knowledge is available, sample size determination can be based on desired levels of power (MacCallum, Browne, and Sugawara, 1996; *see also* STATISTICAL APPROACHES TO DETERMINING SAMPLE SIZES).

Preliminary Data Analysis

One common mistake in applications of SEM seems to be that researchers fail to carefully examine their raw data before calculating the sample covariance matrix on which subsequent analyses are usually based. Space constraints prohibit a detailed discussion of the issues involved (Hair *et al.*, 2006), but researchers should ensure that the data have been coded appropriately, that missing values are dealt with using modern methods such as full information maximum likelihood or Bayesian multiple imputation, that outliers do not distort the sample covariances, and that the necessary distributional assumptions (e.g., multivariate normality) are adequately satisfied.

For the purpose of illustration, we use a dataset collected from 262 female staff members

at two American universities to study the model shown in Figure 2 (see Bagozzi, Baumgartner, and Yi, 1992, for additional information). Respondents completed a questionnaire measuring beliefs, attitudes, and intentions about using coupons for grocery shopping during the upcoming week (on the basis of 7-point scales, except for the second intention item, which was measured on a 11-point scale but linearly transformed to a range of 1–7), and one week later indicated how many coupons from 6 different sources in 21 product categories (plus an "other" category) they had used during the previous week. For simplicity, respondents with missing values were eliminated, leaving an effective sample size of 250, which, combined with a ratio of data points per parameter estimated of about 7, seemed adequate. Not surprisingly, coupon usage was strongly skewed to the right, so a square root transformation was used to normalize the data. Although the data are not multivariate normal (the data are discrete, the univariate skewnesses ranged from −1.35 to 0.83, the univariate kurtoses from −1.43 to 1.91, the relative multivariate kurtosis was 1.17, and tests of univariate and multivariate normality suggested rejection of normality), it appears that the assumption is not too seriously violated. Robustness checks are reported later.

Model Estimation

The goal of estimation is to find values for the five types of unknown parameters, based on the observed covariance matrix, such that the covariance matrix implied by the estimated model parameters is as close as possible to the sample covariance matrix. Model estimation also yields various goodness-of-fit statistics and standard errors for all parameters. A variety of estimation procedures have been proposed, but by far the most frequently used method is maximum likelihood assuming that the data have a multivariate normal distribution. Simulations show that the estimates tend to be robust to violations of normality, but the chi-square test of overall model fit and the estimates of the standard errors may not be. Estimation procedures that do not require multivariate normality are available, but adjustments to the normal-theory-based methods, which are easy to use, seem to work well in practice (see Bentler and Dudgeon 1996, for more details).

Potential estimation problems are nonconvergence (a solution cannot be found within a given number of iterations or within a given time limit) and improper solutions (the values of sample estimates are not possible in the population, such as negative variance estimates). Common causes are poorly specified models, small sample sizes, and few indicators per factor.

There were no estimation problems with the present data set and LISREL 8.80 converged to a proper solution in less than one second (using a normal-theory-based fitting function). Ideally, more than two indicators should be available for all of the constructs, but this was not possible in the present case. Fortunately, the small number of indicators did not result in estimation problems.

Overall Model Evaluation

Before the estimated model is interpreted in detail, researchers should ascertain that the specification is reasonably consistent with the data. If there are serious misspecifications and they are not attended to, the conclusions derived from the model can be seriously misleading. Of course, not all model misspecifications can be detected based on global goodness-of-fit indices, because alternative specifications that may have very different substantive interpretations can be equally consistent with the data (MacCallum *et al.*, 1993). However, if the fit of the model is found to be deficient, it has to be dealt with.

Global fit assessment is based on a summary measure of the discrepancy between the sample and model-implied covariance matrices. Theoretically, the fit of the model can be assessed using a statistic T which, under appropriate assumptions, has a central chi-square distribution under the null hypothesis that the specified model fits perfectly in the population. Depending on the particular assumptions made, there are different T statistics, but ideally they will lead to similar results. On the basis of the likelihood ratio criterion, one compares the likelihood of the hypothesized model to the likelihood of a model with perfect fit (the saturated model) and hopes that T will not be significantly greater than the

number of overidentifying restrictions (i.e., the degrees of freedom of the model).

Unfortunately, there are practical problems with the chi-square test of overall model fit. First, there is evidence that the test is not robust to violations of assumptions such as multivariate normality, although promising corrections to the traditional chi-square test, such as the Satorra–Bentler scaled (robust) test statistic, have been proposed in the literature. Second, the test is based on the accept-support logic, meaning that failure to reject the null hypothesis provides support for the researcher's model. On the one hand, this implies that a model is more likely to be supported when sample size and power are low, even though the chi-square test is only asymptotically valid. On the other hand, since most models are unlikely to be literally true in the population, larger sample sizes will ultimately lead to the rejection of a model even when the misspecification is relatively minor.

Because of these problems, many alternative (mostly descriptive) fit indices have been suggested. They can be classified on the basis of (i) whether they are goodness- or badness-of-fit indices (depending on whether an increase or decrease in the index signals a better fit), (ii) whether they adjudge fit in an absolute or relative sense (stand-alone vs. incremental fit indices, where the model of complete independence of all observed variables is generally used as the baseline model for the incremental indices), (iii) whether they are normed, approximately normed, or nonnormed (i.e., always or usually constrained to fall within a 0–1 interval in sample data, or unconstrained), and (iv) whether or not the fit index imposes a penalty for fitting additional parameters (correction for model parsimony or not). A detailed discussion is beyond the scope of this article. Some researchers question the value of even those indices which have traditionally been recommended in the literature (including the cutoff values associated with their use; see the recent exchange in the May 2007 issue of *Personality and Individual Differences*), but there is some consensus that a few indices based on different conceptual rationales should be used to assess overall model fit, and we briefly mention some of the more promising indices.

An intuitively appealing index is the standardized root mean square residual (SRMR), which summarizes the average size of the standardized residuals. It is a stand-alone badness-of-fit index, normed to fall between 0 and 1, and does not take into account model parsimony. Values up to 0.05 or maybe 0.10 are often considered to reflect satisfactory fit.

Another index is the root mean square error of approximation (RMSEA), which estimates how well the fitted model approximates the population covariance matrix per degree of freedom, using the noncentrality parameter to index error of approximation. It is computed as $\sqrt{(T-df)/((N-1)df)}$, where T is the test statistic, df the degrees of freedom of the model, and N the sample size. A confidence interval for RMSEA is available, which provides information about the precision of the point estimate. RMSEA is a nonnormed, stand-alone badness-of-fit index which imposes a penalty for fitting additional parameters. Models with RMSEA values below 0.05 are assumed to have close fit, and values up to 0.08 or maybe 0.10 are considered acceptable.

The comparative fit index (CFI) is a normed, incremental goodness-of-fit index without a correction for model parsimony. It is also based on the idea of noncentrality and can be computed as $\left[1-\frac{\max[(T_t-df_t),\,0]}{\max[(T_b-df_b),\,(T_t-df_t),\,0]}\right]$, where the subscripts refer to the target and baseline models, respectively. CFI should be at least 0.9 and maybe 0.95 or higher for a well-fitting model.

Finally, the nonnormed fit index (NNFI, the extension of the original Tucker–Lewis index in exploratory factor analysis to SEM) is an approximately normed incremental goodness-of-fit index which penalizes models containing more parameters. It is given by $\left[1-\frac{(T_t-df_t)/((N-1)df_t)}{(T_b-df_b)/((N-1)df_b)}\right]$; the recommended levels are the same as for CFI.

A special class of fit indices are those based on information theory, such as the Akaike information criterion (AIC). They can be used to compare (even non-nested) models and take into account model parsimony. The model with the smallest value of the fit index (which with some definitions may be negative) is selected.

One common problem often observed in the application of many of these fit indices is that, after having to conclude that the specified model does not fit based on the chi-square test, researchers marshal evidence based on some alternative fit indices that the model is a good enough approximation. The goal seems to be to justify the initially proposed structure rather than to learn something new from the data. Furthermore, it is often unclear how much the original model was modified to bring the implied covariance matrix into reasonable congruence with the sample covariance matrix, in which case the overall model evaluation may not be trustworthy because of the dangers of capitalization on chance.

The fit statistics for the illustrative model are as follows. The T statistic based on normal-theory weighted least squares (the minimum of the normal-theory fitting function) was 92.60 (93.63) at 70 degrees of freedom, yielding a p-value of 0.04 (0.03). Thus, the fit based on the conventional chi-square test is borderline at an α-level of 0.05. When the normal-theory-based T was corrected for nonnormality using the Satorra–Bentler scaled test statistic, the resulting value was 83.57 ($p = 0.13$). The SRMR was 0.05, RMSEA 0.036 (with a confidence interval ranging from 0.0096 to 0.054), and both the CFI and NNFI were 0.99. All these indices suggest a very good fit of the model. This is somewhat unusual, but there are several reasons for this result. The model underlying Figure 1 is well established, the assumed belief structure is based on prior research, and the measurement of the remaining constructs is standard. Even when the model seems to fit well, it is advisable to inspect the results in greater detail to evaluate whether modifications are warranted, as described next.

Model Modification

Model modifications are usually motivated by mediocre overall fits of the initially specified model. Two primary tools are used for this purpose: modification indices (Lagrange multiplier tests) and residuals (the differences between the sample and implied covariance matrices). A modification index is reported for each fixed or constrained parameter and it estimates the predicted decrease in the T statistic when a fixed parameter is freely estimated or an equality constraint is relaxed. Associated with each modification index is an expected parameter change (in original and various standardized scale units), which shows the predicted value of the freely estimated parameter. Assuming a chi-square distribution for T, a modification index exceeding 3.84 suggests a significant improvement of the model and a significant parameter estimate when the parameter in question is freely estimated (at an α-level of 0.05).

Residuals show the elements in the observed covariance matrix that are over- or underfitted, and this may also alert the researcher to components of the model that require attention. The raw residuals depend on the scale in which the variables are measured and sampling fluctuations, apart from possible model misspecification, but dividing each residual by the square root of the estimated asymptotic variance corrects for these confounds. These so-called "standardized" residuals (not to be confused with the standardized residuals on which SRMR is based) can be interpreted as z-values and indicate which residuals are larger than expected.

Although modification indices and residuals can be very useful, there are two potential dangers. First, if model modifications are primarily driven by the goal to improve the fit of the model (i.e., additional parameters are added based on the size of the associated modification indices), models that are hard to interpret substantively or theoretically may result. Furthermore, research has shown that data-based model modifications are not always able to recover the "true" underlying structure, and often the values of different modification indices are of similar magnitude, making it difficult to decide which parameter to free. It is therefore crucial that data-based model modifications be tempered by knowledge of the substantive area and theoretical considerations. Second, data-based model modifications may overemphasize idiosyncrasies of the particular data set and therefore might not hold up in future studies (the problem of capitalization on chance). Ideally, data-based model modifications should be reevaluated via more confirmatory follow-up research, or a cross-validation strategy

should be used in which the full data set is split into calibration and validation samples and the generalizability of modifications introduced in the former is examined in the latter.

Some authors (Anderson and Gerbing, 1988) believe that structural equation models should be evaluated in a two-step process (as many as four steps have been proposed), where in the first step no restrictions are imposed on the latent variable model (i.e., all covariances among the latent variables of substantive interest are freely estimated) and attention is focused on assessing the adequacy of the measurement model (e.g., whether indicators load on the "right" constructs, nontarget loadings are small and nonsignificant, the items individually and as sets are sufficiently reliable, and the constructs are discriminant). In the second step, the more constrained structural specification of interest is imposed on the latent variable model and the target model is compared to alternative specifications that are either more or less restrictive than the target model (the endpoints being the null and saturated latent variable models). If this detailed evaluation of the measurement and latent variable models reveals serious problems, appropriate model modifications to either model will be required (see Anderson and Gerbing (1988), for details).

For simplicity, we did not conduct separate analyses for the measurement and structural models in our illustrative example (the measurement model corresponding to Figure 2 had a T statistic based on normal-theory weighted least squares of 62.90 with 63 degrees of freedom). For the model in Figure 2 (which, as can be recalled, had a T statistic of 92.60 with 70 degrees of freedom), there were 11 significant modification indices (out of 131) and 9 significant standardized residuals (out of 105). The largest modification index (12.67) suggested a direct path from rewards to behavior, not mediated by attitudes and intentions. This was also supported by significant positive residuals from the two reward indicators to coupon usage (i.e., the initial model does not fully account for the observed covariance between the two reward indicators and coupon usage). When the path in question was freely estimated ($\gamma_{31} = 0.30$, z-value of 3.6), the T statistic based on normal-theory weighted least squares was 79.21 ($p = 0.19$). Most parameter estimates and z-values in the revised model were very close to those in the initial solution, except that the magnitude of the coefficient from intentions to behavior decreased from 0.49 to 0.41. The variance accounted for in coupon usage increased from 0.34 to 0.37 in the revised model. There were still 10 significant modification indices and several significant negative residuals in the revised model, but the suggested changes were hard to interpret. Therefore, no additional model modifications were considered. Ideally, the added path from rewards to coupon usage should be validated in subsequent research, but since there is evidence from at least one prior study that rewards can influence behavior directly, there is precedence for this modification.

Model Interpretation

A detailed evaluation of the measurement model should involve information about the estimated factor loadings and measurement error (unique) variances (including the variability of the estimates and T-values), evidence about measurement reliability (both for individual items and all indicators of a given construct combined), and some indication that the constructs in the model have discriminant validity (*see also* VALIDITY AND RELIABILITY). Individual-item reliability is simply the squared correlation between a construct ξ_j and an indicator x_i. It can be computed as $\rho_{ii} = \lambda_{ij}^2 \text{ var}(\xi_j)/[\lambda_{ij}^2 \text{ var}(\xi_j) + \theta_{ii}]$. It would be desirable if at least half of the variance of an observed variable were substantive variance rather than measurement error (unique variance), but this is often not the case. Two summary measures of reliability for all indicators of a construct are in common use. Composite reliability is the squared correlation between a construct and an unweighted sum of its indicators. It can be obtained as $\rho_c = (\sum\lambda_{ij})^2 \text{ var}(\xi_j)/[(\sum\lambda_{ij})^2 \text{ var}(\xi_j) + \sum\theta_{ii}]$. Average variance extracted (AVE) is the proportion of the total variance in all indicators of a construct accounted for by the construct, and it is calculated as the average of the individual-item reliabilities. Composite reliability is a generalization of coefficient alpha and therefore the same guidelines apply. Values below 0.6 indicate poor

reliability and above 0.8 are desirable. For AVE, values greater than 0.5 are deemed satisfactory.

Discriminant validity is commonly assessed in the following two ways. First, the correlation between two constructs should be significantly different from unity. This can be tested by constructing a confidence interval around the estimated correlation and checking whether the interval excludes 1. A more stringent test is based on AVE. Specifically, the AVE of the two constructs involved in a correlation should be greater than the squared correlation. This is based on the intuitive notion that a construct should have more in common with its own indicators than with other constructs.

In the illustrative example, all freely estimated loadings were highly significant, the individual-item reliabilities were mostly 0.5 or higher (the two exceptions were the second indicator of rewards with a reliability of 0.48 and the first indicator of encumbrances with a reliability of 0.24). In future research, the latter indicator should probably be modified. The composite reliabilities and AVEs were 0.76 and 0.61 for rewards; 0.88 and 0.78 for inconveniences; 0.70 and 0.44 for encumbrances; 0.88 and 0.65 for attitudes; and 0.92 and 0.85 for intentions. With the possible exception of encumbrances, all constructs seem to have been measured adequately. The largest (disattenuated) correlation is between attitudes and intentions (0.69) and discriminant validity is satisfied, based on both the criteria.

The latent variable model is what researchers will generally be most interested in. The sign, magnitude, and significance of the relationships between the constructs test the hypotheses that the research was designed to investigate. In addition, researchers should report the variance accounted for in each of the endogenous variables so that readers can get an impression of the size of the effects (given that the estimated coefficients often do not have a natural interpretation).

Summary information about the latent variable model in the illustrative example is shown in Table 1. Reward beliefs have a positive influence on attitudes and inconvenience beliefs affect attitudes negatively. Encumbrance beliefs do not have a significant effect on attitudes. More favorable attitudes lead to higher intentions, and more positive intentions encourage greater coupon usage. In addition to the indirect effect of reward beliefs on behavior via attitudes and intentions (point estimate of 0.20, with a standard error of 0.04), there is also a direct effect (point estimate of 0.30, with a standard error of 0.08). Thus, the total effect of 0.50 is composed of a direct effect (61% of the total effect) and an indirect effect (39%). In other words, attitudes and intentions only partially mediate the effect of reward beliefs on coupon usage.

The final model was also analyzed using the Satorra–Bentler robust standard errors for the coefficients in the latent variable model and the results were almost identical. In addition, 100 bootstrap samples were drawn and analyzed and again the results were very similar and the substantive conclusions remained exactly the same.

It should be noted that any causal interpretations implied in the foregoing description are based on theoretical grounds and facilitated

Table 1 Summary information for the latent variable model.

Path	*Parameter*	*Parameter Estimate (Standard Error)*	*Standardized Estimate*	*z-Value*
Rewards → attitudes	γ_{11}	0.44 (0.08)	0.47	5.59
Inconveniences → attitudes	γ_{12}	–0.28 (0.06)	–0.38	–4.85
Encumbrances → attitudes	γ_{13}	–0.04 (0.10)	–0.03	–0.41
Rewards → coupon usage	γ_{31}	0.30 (0.08)	0.24	3.60
Attitudes → intentions	β_{21}	1.09 (0.11)	0.69	9.88
Intentions → coupon usage	β_{32}	0.41 (0.05)	0.48	7.96
R^2(attitudes)	—	0.42	—	—
R^2(intentions)	—	0.48	—	—
R^2(behavior)	—	0.37	—	—

to some extent by the design of the study (*see also* CAUSAL RESEARCH; CONCEPT OF CAUSALITY AND CONDITIONS FOR CAUSALITY). SEM per se does not warrant these conclusions. For example, coupon usage was measured one week *after* respondents indicated their beliefs, attitudes, and intentions. Although the self-report of coupon usage may have conceivably been influenced by people's memory of their previous responses, it is difficult to imagine that coupon usage (measured at a later point) could have influenced the other constructs in the model (measured at an earlier point). On the other hand, beliefs, attitudes, and intentions were all measured in the same questionnaire and the specification of the direction of the effects is based on one well-established theory (the theory of reasoned action), although one could probably adduce other theories to justify different patterns of effects.

Other Uses of SEM and Recent Developments

In this final section we briefly mention some other uses of SEM and recent developments, which cannot be dealt with fully because of space constraints (see the edited book by Hancock and Mueller (2006), for review chapters on many of the topics in this section).

The models described so far are covariance structure models that ignore the means of the observed and latent variables. Furthermore, they are models for a single population. It is possible to extend the models to multiple known populations, in which case the means of the observed and latent variables can be incorporated as well. If latent means and path coefficients are to be compared across different populations, it is necessary to first establish measurement invariance. Depending on the type of comparison to be conducted, different degrees of measurement invariance are necessary (see Steenkamp and Baumgartner (1998), for details). For example, the data in the illustrative application were collected at two different universities and it might be of interest to establish whether the means of the different constructs or the strength of the relationships between constructs are the same for the two groups of respondents.

Traditional multigroup SEM assumes that each observation in the sample can be assigned to a population of interest a priori. Sometimes, however, the researcher may believe that the observations in a sample come from multiple populations, but the population membership of individual observations is not known. In this case, latent variable mixture modeling is applicable. The goal is to determine whether a mixture of multiple populations gave rise to the mean and covariance structure from which the sample under investigation was drawn, and to recover the separate model parameters for each of the multiple populations as well as estimate the mixing proportions. Such analyses may be useful for market segmentation and other investigations of unobserved heterogeneity (Jedidi, Jagpal, and DeSarbo, 1997; *see also* UNOBSERVED HETEROGENEITY).

Structural equation models usually involve only linear relationships (e.g., LISREL stands for linear structural relationships). However, sometimes it is of interest to investigate multiplicative or quadratic relationships of latent variables. For example, theories often specify moderator effects of one variable on the relationships between other variables. If such a theory is to be tested, an interaction model has to be specified. Although moderator hypotheses can be tested with multigroup structural equation models when the moderator is binary or the number of levels of the moderator is small, a different approach is needed when the moderator is continuous. Several approaches have been proposed to specify nonlinear effects in SEM. One key advantage of these models is that measurement error, which has even more damaging effects when interactions and quadratic effects are present, is taken into account explicitly. However, these models are also quite complex and they frequently require assumptions that might offset their perceived advantages.

The basic structural equation model assumes that the data are a simple random sample from a single underlying population. However, complex survey designs are sometimes used, where observations are sampled in multiple stages (e.g., first companies are sampled, and then salespeople within companies), the total population is stratified (e.g., by male and female

salespeople), and selection probabilities are unequal across strata (e.g., female salespeople have a higher probability of being included in the sample; *see also* SAMPLING TECHNIQUES; PROBABILITY SAMPLING). Multilevel SEM and sampling weights can be used to deal with these complications. For example, a researcher could specify a multilevel structural equation model and study (i) whether more motivated salespeople in a company are more likely to attain higher sales (within-company analysis) and (ii) whether companies whose salespeople are more motivated on average tend to have higher sales (between-company analysis).

Finally, SEM can be used for longitudinal analyses in which a researcher wants to study the trajectory of change in some variable(s) of interest and explain various aspects of the change trajectory (e.g., the linear rate of change) based on other variables, or use the change trajectory as an antecedent variable. In such an analysis, which is referred to as *latent curve* or *latent growth modeling*, the latent variables indicated by the repeated measurements (e.g., amount of cola consumed per year over a number of years) are individual-level curve parameters (e.g., intercepts and slopes if a linear trajectory is assumed, although other specifications are possible), which can subsequently serve as endogenous or exogenous factors in more detailed investigations of the change process.

Conclusion

SEM has become a valuable addition to the methodological toolbox of researchers in social sciences in general and marketing in particular. It combines a concern with measurement, which takes into account the inherent fallibility of single indicators of constructs, with the opportunity to model complex patterns of relationships among constructs. The scope of SEM is constantly being expanded and it can be expected that SEM will continue to flourish in empirical research in marketing.

Bibliography

Anderson, J.C. and Gerbing, D.W. (1988) Structural equation modeling in practice: a review and recommended two-step approach. *Psychological Bulletin*, **103**, 411–423.

Bagozzi, R.P. and Baumgartner, H. (1994) The evaluation of structural equation models and hypothesis testing, in *Principles of Marketing Research* (ed R.P. Bagozzi), Blackwell Publishers, Cambridge, pp. 386–422.

Bagozzi, R.P., Baumgartner, H., and Yi, Y. (1992) State versus action orientation and the theory of reasoned action: an application to coupon usage. *Journal of Consumer Research*, **18**, 505–518.

Bandalos, D.L. and Finney, S.J. (2001) Item parceling issues in structural equation modeling, in *New Developments and Techniques in Structural Equation Modeling* (eds G.A. Marcoulides and R.E. Schumacker), Lawrence Erlbaum, Mahwah, pp. 269–296.

Baumgartner, H. and Homburg, C. (1996) Applications of structural equation modeling in marketing and consumer research: a review. *International Journal of Research in Marketing*, **13**, 139–161.

Bentler, P.M. and Dudgeon, P. (1996) Covariance structure analysis: statistical practice, theory, and directions. *Annual Review of Psychology*, **47**, 563–592.

Bollen, K.A. (1989) *Structural Equations with Latent Variables*, John Wiley & Sons, Inc., New York.

Diamantopoulos, A., Riefler, P., and Roth, K.P. (2008) Advancing formative measurement models. *Journal of Business Research*, **61**, 1203–1218.

Hair, J.F. Jr, Black, W.C., Babin, B.J. *et al*. (2006) *Multivariate Data Analysis*, 6th edn, Pearson Prentice Hall, Upper Saddle River.

Hancock, G.R. and Mueller, R.O. (eds) (2006) *Structural Equation Modeling: A Second Course*, Information Age Publishing, Greenwich.

Jedidi, K., Jagpal, H.S., and DeSarbo, W.S. (1997) Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science*, **16**, 39–59.

MacCallum, R.C., Browne, M.W., and Sugawara, H.M. (1996) Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, **1**, 130–149.

MacCallum, R.C., Wegener, D.T., Uchino, B.N., and Fabrigar, L.R. (1993) The problem of equivalent models in applications of covariance structure models. *Psychological Bulletin*, **114**, 185–199.

MacKenzie, S.B., Lutz, R.J., and Belch, G.E. (1986) The role of attitude toward the ad as a mediator of advertising effectiveness: a test of competing explanations. *Journal of Marketing Research*, **23**, 130–143.

Steenkamp, J.-B.E.M. and Baumgartner, H. (1998) Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, **25**, 78–90.

univariate techniques

Rex Yuxing Du

Generally speaking, all statistical techniques can be characterized as either univariate or multivariate. Univariate techniques should be used for analyzing data when there is a single measurement (variable) concerning each element in the sample or when there are several measurements on each element but each measurement is analyzed independently or in isolation. On the other hand, multivariate techniques are applicable for data analyses when there are two or more measurements of each element and these measurements are examined *simultaneously*. The key distinction between univariate techniques and multivariate techniques lies in their different focuses. Univariate techniques are concerned with the levels (means) and distributions (variances) of each individual variable. For example, a sample of respondents may rate a brand at two different times. To determine whether there is a significant shift in the respondents' ratings over time, one can use a univariate technique, focused on examining the mean and variance of the differences in the respondents' ratings. By contrast, multivariate techniques are concerned with the degree of dependence or interdependence among two or more variables (correlations or covariances). To continue the previous example, suppose one has also collected data on the respondents' prior brand usage, ad exposure, category purchases, and so on, multivariate techniques can be used to examine how these variables covary along with brand ratings. This article focuses on a few univariate techniques most frequently encountered in marketing research. The analysis of frequencies, which does not involve statistical testing, is discussed in FREQUENCY DISTRIBUTION. For more details on various commonly used multivariate techniques, *see* CLUSTER ANALYSIS, CONJOINT ANALYSIS, EXPLORATORY FACTOR ANALYSIS, MODELS FOR CATEGORICAL DEPENDENT VARIABLES, MULTIDIMENSIONAL SCALING OF PREFERENCE DATA, DISCRIMINANT ANALYSIS FOR MARKETING RESEARCH APPLICATIONS, STRUCTURAL EQUATION MODELING, ANALYSIS OF VARIANCE AND COVARIANCE, MULTIPLE REGRESSION.

Different univariate techniques can be classified on the basis of whether the focal variable is metric or nonmetric. A metric variable is measured on an interval or a ratio scale (also *see* PRIMARY SCALES OF MEASUREMENT). A satisfaction rating based on a five-point scale is an example of metric variable. Nonmetric variable is measured on a nominal or an ordinal scale. Survey questions with a yes or no response offer data in a nonmetric format.

For both nonmetric and metric variables, univariate techniques can be further classified based on whether a single sample or multiple samples are involved. Note that the number of samples is determined on the basis of how the data are treated for the purpose of analysis, as opposed to on the basis of how the data were gathered. For example, data for the general population may have been collected as a single sample, but if the analysis entails examining the potential differences between customers and noncustomers, two-sample univariate techniques should be applied.

Furthermore, when multiple samples are involved, the appropriateness of a univariate technique depends on whether the samples are independent or dependent (often also referred to as *paired*). Samples are deemed as independent if observations can be thought of as randomly drawn from different populations and there is no reason to believe any pair of observations (from different samples) are somehow related to each other. For example, to examine gender differences, data on males and females can be treated as from independent samples. On the other hand, two samples are deemed as dependent when each observation from one sample can somehow be paired with another observation from the other sample. For example, two observations are paired if they come from before-and-after samples of the same individual or right-side-and-left-side samples of the same individual. It is important to keep in mind that it is not the before-after that makes two observations paired, but the fact that they come from the same individual.

In sum, Figure 1 presents a classification of univariate techniques on the basis of (i) whether the focal variable is metric or nonmetric, (ii)

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

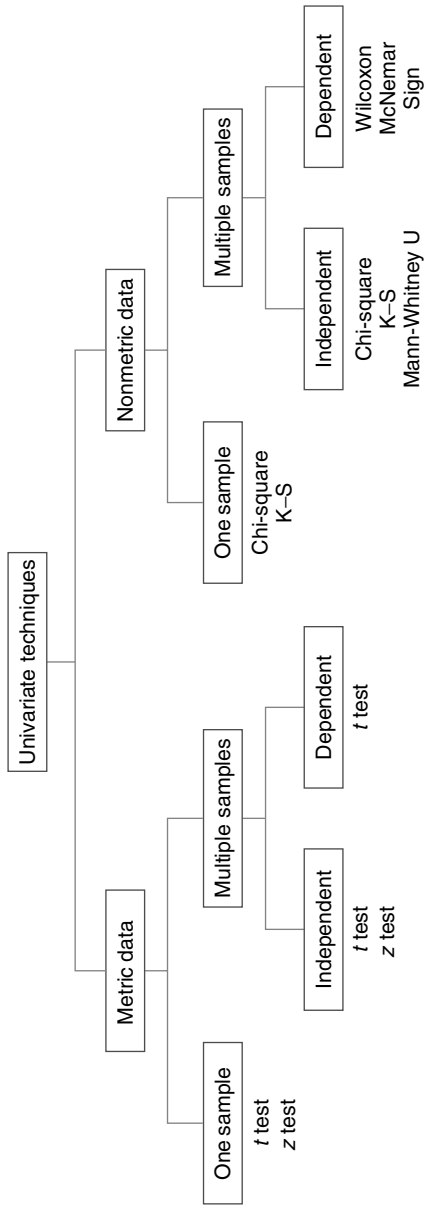


Figure 1 A classification of univariate techniques (adapted from Aaker, Kumar, and Day (2006). © John Wiley & Sons Inc., 2006).

whether there is one or multiple samples, and (iii) in the case of multiple samples, whether they are independent or dependent. The rest of the article focuses on three most frequently used univariate techniques: *t*-tests with metric data (interval and ratio), chi-square tests with nominal data, and Kolmogorov–Smirnov tests (K-S tests) with ordinal data. All three techniques can be applied to data from one or multiple samples. As for the other less frequently used and less versatile techniques, we provide a brief description on when they are appropriate and leave the details for more technical texts.

T-TESTS WITH METRIC DATA

One sample t-test. In marketing research, the researcher is often interested in comparing the mean of a single variable against a known or a given standard. For example, in doing research for a restaurant, the researcher might be interested in knowing whether patrons are satisfied with the restaurant's food. Respondents have provided their answers based on a seven-point interval scale where 1 = "Very Dissatisfied" and 7 = "Very Satisfied." Before attempting to address the "whether customers are satisfied" question, the researcher must perform a couple of tasks. First, he/she must develop the null and alternative hypotheses. For example, the researcher may be interested in determining whether the mean satisfaction level of the population exceeds 4, the neutral value on a seven-point scale. In that case, the null hypothesis may be stated as "the mean satisfaction level of the population is no greater than 4," and the alternative hypothesis as "the mean satisfaction level of the population is greater than 4." After developing the hypotheses, the researcher needs to select the level of significance for rejecting the null hypothesis (say, $\alpha = 0.05$, which means the chances of rejecting the null hypothesis when it is in fact true needs to be smaller than 5%).

After formulating the null and alternative hypotheses and selecting the significance level, the researcher can conduct a one-sample *t-test*, which is commonly used to provide inferences for making statements about the sample mean of a single metric variable. The *t*-test is based on the Student's *t* statistic. The *t* statistic assumes that the variable (X) is normally distributed across the population, with the mean of the population (μ) known or assumed to be known, and the standard deviation of the population (σ) estimated from the sample. If these assumptions hold, then we know that the values of $t = (\overline{X} - \mu/s)$ would be *t* distributed[1] with $n - 1$ degrees of freedom, where $\overline{X}$ is the sample mean, s the sample standard deviation, and n the sample size. The researcher can reject the null hypothesis and state that the mean satisfaction level of the population is above neutral if the value of the calculated *t* statistic (based on the sample) exceeds the critical *t* value (one-tailed in this case) for $n - 1$ degrees of freedom and a significance level of α. For more details on hypothesis testing related to differences, *see* HYPOTHESIS TESTING RELATED TO DIFFERENCES – PARAMETRIC TESTS.

Note that if the population variance (σ), like mean (μ), is assumed to be known, as opposed to estimated from the sample, a z test would be appropriate, where the value of $z = (\overline{X} - \mu/\sigma)$ would be distributed standard normal. However, because the *t* distribution approaches the standard normal distribution as the degrees of freedom increases, for large samples (say, 120 or more), the *t* test, and the z test are, for all practical purposes, equivalent.

Two independent samples t-test. Instead of simply making statements about a single population mean against a known or a given standard, which requires only one sample, marketing researchers are often interested in drawing conclusions about means from two different populations, which would require two independent samples; for example, users versus nonusers of a brand in terms of their category requirements, or high-income versus low-income consumers in terms of their time spent watching television. In such cases, *t*-test would again be appropriate.

After drawing independent samples of sizes n_1 and n_2 from two different populations and computing the means ($\overline{X}_1$ and $\overline{X}_2$) and standard deviations (s_1 and s_2), the researcher can compute the *t*-statistic as $t = (\overline{X}_1 - \overline{X}_2 - \Delta/\sqrt{s_1^2/n_1 + s_2^2/n_2})$, where Δ is the hypothesized difference between the population means (0 if trying to determine whether the means are equal). The next step

is to look up the critical t value in the t-table, with the number of degrees of freedom being the smaller of $n_1 - 1$ and $n_2 - 1$, and the significance level determined by the researcher. Whether to reject the null hypothesis depends on whether the computed t-statistic exceeds the critical value from the t-table.

If it is appropriate to assume that the two populations have equal (but unknown) variance (and therefore the same standard deviation), a pooled variance estimate can be computed from the two sample standard deviations as $s_P^2 = ((n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2 / n_1 + n_2 - 2)$. The formula for computing the t-statistic using the pooled variance is $t = (\overline{X}_1 - \overline{X}_2 - \Delta / \sqrt{s_p^2\,(1/n_1 + 1/n_2)})$, and the corresponding number of degrees of freedom is $n_1 + n_2 - 2$.

Note that using pooled variance in a t-test is, in general, more likely to yield significant results than using separate variances. To determine whether the variances of the two populations are equal, an F-test may be performed, with the F-statistic computed as $F = (s_1^2/s_2^2)$. The critical F value for a given significance level can be found in an F-table, with two sets of degrees of freedom: $n_1 - 1$ based on the numerator and $n_2 - 1$ based on the denominator. If the computed F-statistic exceeds the critical F value, the null hypothesis of equal variance needs to be rejected and t-test based on separate variances should be used; otherwise, the t-test based on pooled variance can be used.

Two dependent samples t-test. In many marketing research applications, two related (or paired) observations are gathered from the same respondents. For example, a sample of respondents may indicate the relative importance of two attributes of a product, or rate two competing brands in one survey, or rate one brand at two different times (e.g., before and after an advertisement campaign). In these cases, it is the *difference* between the two related observations that is examined, which can be achieved by a paired-samples t-test. To compute the t-statistic for two dependent samples, the difference between the related observations, denoted by D, is formed and its mean ($\overline{D}$) and standard deviation (s_D) computed. The relevant formula is $t = (\overline{D} - \mu_D / s_D)$, with the degrees of freedom being $n - 1$, where n is the number of pairs and μ_D the hypothesized mean of the differences.

Chi-square Tests with Nominal Data

Many variables in marketing research are measured on nominal scales. With these variables, the researcher is often interested in addressing questions such as "Do the numbers of responses that fall into different categories differ from what is expected?" or "Does the pattern of observed frequency counts based on one sample differ from those based on other samples drawn from different populations?" the *Chi-square test*, which is often referred to as the "*goodness-of-fit*" test, is a commonly used univariate technique to answer these questions. Below are two specific examples of research questions that could be examined by using chi-square tests.

- Do we have the same proportions of customers who visit our stores on an infrequent, moderately frequent, or very frequent basis?
- Does usage of mobile phone – never, low, moderate, and high – differ between the male population and the female population (two independent samples)?

The first question involves one sample and the null hypothesis is that one-third of the customer base falls into each of the three categories. The second question involves two independent samples and the null hypothesis is that the percentages of male mobile phone users falling into each of the four categories follow the same pattern as those of female users. In both cases, the general formula for calculating the *chi-square statistic* is $\chi^2 = \sum_{i=1}^{K} ((O_i - E_i)^2 / E_i)$, where K = number of mutually exclusive subgroups of respondents (differing in terms of either the categories of their responses or the populations to which they belong), O_i = number of respondents in subgroup i, and E_i = number of respondents expected in subgroup i if the null hypothesis is true. Under the null hypothesis, χ^2 follows the chi-square distribution, which is completely determined by its degrees of freedom, with the mean being

equal to the number of degrees of freedom and the variance equal to two times of the mean. When the degrees of freedom is large, the chi-square distribution is approximately normally distributed.

In the case of one sample, $E_i = p_i \times n$, with $n =$ the total number of respondents and $p_i =$ percentage of respondents falling into category i according to the null hypothesis. The degrees of freedom for one-sample chi-square test is $K - 1$. In the case of two independent samples, for subgroup i that belongs to sample s and response category r $E_i = (n_s n_c / n)$, where $n_s =$ total number of respondents from sample s, $n_c =$ total number of respondents falling into category c, and $n =$ total number of respondents. The degrees of freedom for two-sample chi-square test is $(2 - 1) \times (C - 1) = C - 1$, with C being the total number of response categories.

Kolmogorov–Smirnov (K–S) Tests with Ordinal Data

In many marketing research applications, the variable of interest is measured on an ordinal scale (e.g., never, infrequent, moderately frequent, or very frequent), where the spacing between adjacent scale points cannot be assumed to be the same. For convenience, such variables can be treated as if they were nominal, which, however, would be inefficient use of data because the scale points do follow a rank order. To fully utilize such information, K–S tests are often preferred. The K–S test compares the observed cumulative proportion with an alternative distribution. In the case of data from one sample, the alternative distribution is specified according to the null hypothesis. In the case of data from two independent samples, the alternative distribution is given by data from the other sample.

The K–S test is based on the maximum value of the absolute difference between O_i (the observed sample cumulative proportion through the ith ranked category) and A_i (the corresponding value given by the null hypothesis). The test statistic is $K = Max\,|A_i - O_i|$. Whether to reject the null hypothesis depends on the value of K; the larger the K is, the more confidence one should have in rejecting the null hypothesis. For significance level of 0.05, the critical value of K for large samples (over 35) is given by $1.36/\sqrt{n}$ (Kanji, 1999).

Note that the K–S test examines whether two distributions are the same, taking into account any differences between the two distributions, including the median, dispersion, and skewness. If the objective is to determine whether two independent samples are drawn from populations with the same median, *two-sample median test* would be more appropriate. If the objective is to determine whether there is a significant difference in the location (or central tendency) of two independent populations, the *Mann–Whitney U test* can be used, which corresponds to the two independent samples t-test for metric data with a pooled variance estimate.

Wilcoxon matched-pairs signed-ranks test (or simply *Wilcoxon test*) is an important test for examining differences in the central tendency of two samples with paired observations, which takes into account the magnitude of the differences measured on an ordinal scale. It first computes the differences between the pairs of observations and ranks the absolute differences. The next step is to sum the positive and negative ranks. The test statistic, z, is computed from the positive and negative rank sums. Under the null hypothesis of no difference, z is distributed standard normal with mean zero and standard deviation one for large samples. This test corresponds to the paired samples t-test discussed earlier for metric data (Malhotra, 2006, p. 488).

Another paired samples test for data collected by ordinal scales is the *sign test*. This test is not as powerful as the Wilcoxon test as it only compares the signs of the differences between pairs of variables without taking into account the magnitude of the differences. In the special case of a binary variable where the researcher wishes to test differences in proportions, the *McNemar* test can be used. Alternatively, the chi-square test can also be used for binary variables. For more on nonparametric tests, *see* NONPARAMETRIC TEST and standard statistical literature such as Higgins (2002) and Pett (1997).

Endnotes

[1] The t distribution is similar to the standard normal distribution in appearance, both being bell-shaped and symmetric. Relative to the standard normal distribution, however, the t distribution has less area in the center and more in the tails (i.e., more variance), which is, intuitively, owing to the fact that population standard deviation σ is unknown and is estimated by the sample standard deviation s, which itself varies from sample to sample. As a result, one must go a larger number of standard deviations from zero to encompass a certain percentage of values from the t distribution than is the case with the standard normal distribution.

Bibliography

Aaker, D.A., Kumar, V., and Day, G.S. (2006) *Marketing Research*, John Wiley & Sons, Inc., Hoboken.

Higgins, J.J. (2002) *Introduction to Modern Nonparametric Statistics*, Duxbury, Pacific Grove.

Kanji, G.K. (1999) *100 Statistical Tests: New Edition*, Sage Publications, Thousand Oaks.

Malhotra, N.K. (2006) *Marketing Research: An Applied Orientation*, Pearson Prentice Hall, Upper Saddle River.

Pett, M.A. (1997) *Nonparametric Statistics for Health Care Research*, Sage Publications, Thousand Oaks.

common methods bias

Wagner A. Kamakura

Assume a market researcher is interested in studying the relationship between customer satisfaction and customer loyalty. On the basis of theoretical considerations, one would expect that customer satisfaction measures would be correlated with measures of loyalty. However, if the measures of customer satisfaction and the measures of loyalty share common methods, those methods may produce a systematic effect on the observed correlation between the measures. In other words, subjects answering questions about satisfaction and loyalty within the same survey would be likely to correlate their answers to the loyalty questions highly with their answers to the customer satisfaction questions, simply to avoid cognitive dissonance. Therefore, the survey does not gather the "true" satisfaction and loyalty scores from the subjects because of common methods bias, which is likely to produce a spurious correlation between the two scores measuring customer satisfaction and loyalty.

Method biases are a problem because they are one of the main sources of measurement error. A measurement error threatens the validity of conclusions drawn from observed associations among measures and is known to have both a random and a systematic component. Although, both types of measurement errors are problematic, systematic measurement error is more critical because it produces spurious correlations that are due to the measurement instruments, rather than to the true association among the constructs being measured.

The influence of common methods variance has become a major concern in survey-based research (Podsakoff *et al.*, 2003), particularly for research involving self-reported measures (Spector, 2006). Much of the evidence of the extent to which method variance is present in behavioral research comes from meta-analyses of multitrait–multimethod studies (Campbel and Fiske, 1959; Cote and Buckley, 1987; Williams, Cote, and Buckley, 1989; Doty and Glick, 1998), which found that approximately one-quarter of the variance in survey-based research might be due to systematic sources of measurement error like common method biases (Cote and Buckley, 1987). Common methods variance owing to systematic measurement error can cause observed correlations among variables to differ from their "true" population values (Doty and Glick, 1998), a manifestation of common methods bias. Common methods variance is not the same as common methods bias (Spector, 2006); whereas the former implies that variance in observed scores is partially attributable to a methods effect, common methods bias refers to the degree to which correlations are altered (usually inflated) owing to a methods effect.

Podsakoff *et al.* (2003) summarized the literature on common methods variance, identifying four main sources for this major threat to research validity as follows:

- common rater (e.g., SOCIAL DESIRABILITY BIAS, leniency)
- item characteristic effects (e.g., item ambiguity; *see* ITEMIZED RATING SCALES (LIKERT, SEMANTIC DIFFERENTIAL, AND STAPEL); PRIMARY SCALES OF MEASUREMENT)
- item context effects (e.g., priming effects, grouping of items)
- measurement context effects (e.g., simultaneous measurement of predictor and criterion variables).

Common methods bias is typically assessed and controlled for via CONFIRMATORY FACTOR ANALYSIS. One of the available approaches (see Podsakoff *et al.*, 2003 for a detailed discussion) is to specify each item in the measurement scale not only as an indicator of its substantive construct but also of a latent method factor, thereby controlling for common methods variance through the latent methods factor. Another popular approach is to use confirmatory factor analysis to directly model Campbell and Fiske's (1959) multitrait–multimethod matrix, specifying both latent trait and latent method factors (Doty and Glick, 1998).

Bibliography

Campbell, D.T. and Fiske, D.W. (1959) Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, **56** (2), 81–105.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Cote, J.A. and Buckley, M.R. (1987) Estimating trait, method, and error variance: generalizing across 70 construct validation studies. *Journal of Marketing Research*, **24** (3), 315–318.

Crampton, S.M. and Wagner, J.A. (1994) Percept percept inflation in microorganizational research: an investigation of prevalence and effect. *Journal of Applied Psychology*, **79**, 67–76.

Doty, D.H. and Glick, W.H. (1998) Common methods bias: does common methods variance really bias results? *Organizational Research Methods*, **1**, 374–406.

Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y., and Podsakoff, N.P. (2003) Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology*, **88** (5), 879–903.

Spector, P.E. (2006) Method variance in organizational research: truth or urban legend? *Organizational Research Methods*, **9**, 221–232.

Williams, L.J., Cote, J.A., and Buckley, M.R. (1989) Lack of method variance in self-reported affect and perceptions at work: reality or artifact? *Journal of Applied Psychology*, **74**, 462–468.

criteria for evaluating secondary data

Frank J. Mulhern

Secondary data are market research data collected for a purpose other than the one on hand. The data have the advantage of being much less expensive and more quickly available than primary data. However, because secondary data were collected for a different purpose, one must carefully evaluate their appropriateness for a given market research situation. Secondary data should be evaluated with the following criteria in mind:

- *Accuracy*: The accuracy of secondary data refers to the intrinsic quality of the data. That is, the extent to which one can assume the reported data are correct. Problems with accuracy can result from several factors including nonresponse in surveys, missing values in a database, data entry mistakes, and tabulation or reporting errors. Concern should also be placed on the reliability and validity of the data (*see* RESEARCH RELIABILITY AND VALIDITY).
- *Relevance*: The relevance of secondary data represents their usefulness for the purpose at hand. Even if data are highly accurate, they may be of little value if they lack relevance for the given research situation. Relevance stems from several aspects including the following:
  - *Method of data collection*: The data collection method has enormous effect on many aspects of the data and their relevance for a given purpose. In particular, one should determine if the data came from self-reports, as in the case of survey data (*see* SURVEY RESEARCH) versus behavioral measures such as purchase histories (*see* INTERNAL DATABASES FOR MARKETING RESEARCH).
  - *Units of measure*: Using secondary data precludes the ability to specify the units of measure. For example, a researcher wanting to analyze market share of dollar sales may be unable to do so with data on market share of unit volume.
  - *Level of aggregation*: Users of secondary data must work with levels of aggregation that exist in the data. For example, brand sales for a liquid product might be expressed in terms of the number of packages of all sizes, the equivalence of one size such as quart bottles, or total volume. Aggregation of consumers is also an issue. Data can be reported at different levels of aggregation such as individual consumers, households, or various levels of geographic market areas.
  - *Time increments*: Many forms of data involve time increments such as days, weeks, months, or quarters. While small time increments can be aggregated into larger ones (e.g., daily data can be combined into weekly), larger time units cannot be disaggregated into smaller ones.
  - *Data format*: Working with secondary data can be difficult because data often arrive in undesirable formats that require substantial preparation before analysis can take place. Some datasets come in formats that can only be accessed by certain software.
- *Data documentation*: Secondary data should be accompanied by information on the variable definition, method of data collection, and the like. Researchers reporting on analyses of secondary data will be expected to explain the definition and interpretation of all the variables in the dataset.
- *Currency and timeliness*: Currency refers to how recent a time period the data represent. Recently collected data are generally more desirable. Timeliness refers to how quickly the data can be obtained and when future updates may be available. A common advantage of secondary data is that they can be obtained much more quickly than primary data.
- *Cost*: Secondary data can be expensive to purchase. In addition to the upfront costs of acquisition, one should consider additional costs for data preparation, storage, and so on.
- *Usage terms*: Many secondary data sources place restrictions on who may use the data, what the data may be used for, and the length of time a user may access them. Users of secondary data should be aware of such restrictions.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

2 criteria for evaluating secondary data

By evaluating prospective sources of secondary data with these criteria, market researchers can take advantage of the convenience and cost-effectiveness of secondary data (Kiecolt and Nathan, 1985).

Bibliography

Kiecolt, K.J. and Nathan, L.E. (1985) *Secondary Analysis of Survey Data*, Sage Publications.

internal databases for marketing research

Frank J. Mulhern

Direct marketing and database marketing are two highly related practices that combine the use of customer-level information with marketing communications through direct media. Direct marketing refers to any marketing communication that uses a direct media technology to communicate to individuals, in contrast to mass media, which communicates to large audiences. Direct media include mail, telephone (including mobile), and print catalogs as well as interactive forms of media such as e-mail and online advertising. Database marketing refers to the practice of gathering, compiling, and analyzing information on individual customer purchase behavior and using that information to select target consumers and execute communications. Direct marketing is characterized by the ability to address individual consumers by name, target consumers based on knowledge of past purchases, communicate with direct response mechanisms that facilitate purchasing, and measure consumer response with precision.

Direct marketing is most effective when it is based on detailed information on consumer behavior stored in customer databases. Customer databases are constructed with information on purchase behavior obtained from a variety of sources including retail stores, response to previous direct marketing, online shopping, and media subscriptions. Databases also may contain descriptive information on individuals obtained from public sources such as home purchases and car registrations.

In contrast to survey research, which is self-reported, customer databases represent a form of unobtrusive data that reflects actual behavior. Marketers analyze the information in databases to identify target customers and formulate communications strategies. Customers are often targeted on the basis of their estimated value to a firm. Customer value represents the past or future value of a customer on the basis of the revenue a customer provides through purchases and the costs incurred in serving that customer. A common approach to measuring customer value is to compute three elements: *recency* – how much time has passed since a previous purchase, *frequency* – how frequently a person has made a purchase, and *monetary value* – the total dollar value of past purchases. Database marketers use scoring models that combine such information into a single score to prioritize who should be targeted. Direct marketing communications are then used to solicit responses from the consumers with the highest scores.

A second area of database analysis is customer segmentation. A major advantage of customer databases is they allow for precise specification of market segments compared to the more limited information used in segmentation with demographic or survey data (*see* SURVEY RESEARCH). Database marketing allows for precise segmentation based on past purchases, price and promotion responsiveness, and projected lifetime value. While the same multivariate statistical methods used for survey research can be used for database segmentation, additional classes of analysis, known as *data mining*, can also be used. Data mining consists of methods that systematically analyze large databases to identify patterns or provide summary information.

A third area of database analysis is customer response analysis. These are similar to market response models that assess the impact of marketing mix variables on buying behavior. However customer response analysis is more precise because it reveals exactly which customers respond to which marketing communications.

As media becomes more digital, the direct, and database marketing approaches are becoming more central to marketing as a whole. Mass marketing approaches pioneered in the print and broadcast era are giving way to more precise and efficient marketing communications driven by information in customer databases and the capabilities of direct marketing and database marketing. As television becomes digital and merges with the Internet, the direct and database marketing model will dominate advertising as real-time information on audience television viewing is captured in databases and used to serve targeted ads. In many ways, direct marketing and database marketing represent a revolution in marketing practices that feature a high level of information and the ability to communicate with individual consumers.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

A few key limitations challenge the practice of direct marketing and database marketing. Direct media can be very bothersome to consumers as unwanted telephone calls, mail, e-mail, and online ads intrude into people's lives. Through technologies and other means, consumers attempt to avoid the intrusiveness of direct marketing. Privacy concerns exist because many consumers oppose the collection and use of extensive information on their purchase behavior. Many countries have extensive regulation of what marketers can or cannot do with information in customer databases. A common practice, particularly for online and mobile communications, is to limit direct marketing to consumers who opt-in to solicitations from specific companies. Finally, the development of new communications technologies forces direct marketers to continually modify practices to accommodate new business models and communication practices. Direct marketers are adjusting their approach to accommodate a world where consumers, not marketers, have control over media experiences.

Bibliography

Hughes, A.M. (2005) *Strategic Database Marketing: The Masterplan for Starting and Managing a Profitable, Customer-Based Marketing Program*, 3rd edn, McGraw-Hill, New York.

Shepard, D. (1999) *The New Direct Marketing: How to Implement a Profit-Driven Direct Marketing Strategy*, 3rd edn, McGraw-Hill, New York.

Stone, B. and Jacobs, R. (2008) *Successful Direct Marketing Methods*, 8th edn, McGraw-Hill, New York.

ethnograpic research

Linda L. Price

Ethnography is a research practice that places researchers in the midst of what they study. That is, it takes place "in the field" rather than in experimental setups or highly structured interview situations (*see* FIELD WORK/DATA COLLECTION PROCESS). Ethnography involves participating overtly or covertly in people's daily lives for an extended period of time, watching what happens, listening to what is said, asking questions, and gathering whatever data are available to inform the inquiry.

At its core, ethnography attempts to describe and interpret, cultural and social life (*see* SUBCULTURES). It has been defined as the science of cultural description. Terms such as "*total immersion*," "*thick description*," "*deep hanging out*," and "*subjective soaking*" used to describe ethnography, convey that it is directed at obtaining *and* representing a deep understanding of the lived experience of people as it unfolds in a particular cultural context. Ethnography refers both to the fieldwork and the representations based on that fieldwork. Most ethnographic research is concerned with producing descriptions and explanations of particular phenomena or with developing theory rather than with testing a set of hypotheses. The ethnographer is interested in depth and understanding rather than breadth and prediction.

Ethnographers typically draw on a range of sources of data, but participant and nonparticipant observation, interviews, conversations, and informant diaries are among the most common types of data used (*see* OBSERVATION METHODS). Data collection is relatively unstructured. The research design is a continuous reflexive process that operates throughout each stage of a project. Ethnographers anticipate that the initial questions that motivated the research will be refined or even transformed over the course of the investigation. Categories used for interpreting what people say or do are not typically built into the data collection process, as would be the case with many other types of data collection such as observation schedules, experiments, or questionnaires, but are instead generated out of the process of data analysis. The focus is generally small scale to facilitate in-depth study, but multisited and multivocal ethnographies are increasingly common especially for ethnographers doing commercial market research. Gaining access to the data can be particularly problematic in ethnography, because the researcher often operates in settings where the researcher has little power. New tools such as netnography, used in on-line social systems, and consumer-generated video diaries have expanded ethnographer access to naturally unfolding consumer and marketer actions and practices (*see* ONLINE CONSUMPTION).

Analysis of data is open-ended, flexible, and adaptive. No two ethnographies have ever been conducted in exactly the same way. Ethnographers attempt to understand the informants' points of view (emic) and to portray broader cultural meanings (etic). Data analysis involves interpreting meanings, functions, and consequences of actions and practices of people and institutions. Systematic reading and viewing of collected data and field notes are used to ground tentative themes and hypotheses to the data and generate new ones. Researchers often undertake a process termed *open coding* (unrestricted line by line coding) to identify and extract themes and issues. Next, ethnographers begin to notice and create records of patterns in the conversations and activities of people depicted in field notes. Several commercial programs are available for qualitative data storage, handling and retrieval, but these programs do not analyze data in a way comparable to quantitative software. In reporting themes and patterns observed in the data, ethnographers may rely on inductive content analysis to suggest the magnitude or frequency of identified themes and patterns or alternatively compose an ethnographic narrative account connecting what people say and do with available theoretical interpretations (*see* CONTENT ANALYSIS). The latter approach is most common in consumer and marketing ethnographies. A number of different criteria might be used to evaluate the trustworthiness of the account including adequacy, resonance, integration, innovation, and usefulness.

Ethnography is perhaps the oldest methodology of social investigation. Over the past two decades marketing, and consumer theory

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

and research have established ethnography as an important marketing research tool. Ethnography is useful for a wide range of problems including understanding disjunctures between what people say versus what people do, identifying taken for granted practices and linking them to shared cultural templates and meanings, and uncovering complex interplays of marketing phenomena and people in social systems. Ethnography, or quasi-ethnography, is now a standard offering in the world of commercial market research. The conduct of ethnography can seem deceptively simple, but it is a demanding activity, requiring diverse skills, including dealing with uncertainty and the ability to identify patterns in complex arrays of social meanings and behaviors, and link these patterns to existing social theories.

Bibliography

Arnould, E.J. and Wallendorf, M. (1994) Market-oriented ethnography: interpretation building and marketing strategy formulation. *Journal of Marketing Research*, **31**, 484–504.

Atkinson, P. and Hammersley, M. (2006) *Ethnography: Principles in Practice*, 3rd edn, Routledge Press, New York.

Belk, R.W. (2006) *Handbook of Qualitative Research Methods in Marketing*, Edward Elgar Publishing, Inc., Northampton.

Berg, B.L. (1998) *Qualitative Research Methods in the Social Sciences*, 3rd edn, Allyn & Bacon, Needham Heights.

Kozinets, R.V. (2002) The field behind the screen: using netnography for marketing research in online communities. *Journal of Marketing Research*, **39** (1), 61–72.

Mariampolski, H. (2005) *Ethnography for Marketers: A Guide to Consumer Immersion*, Sage Publications, Thousand Oaks.

Spiggle, S. (1994) Analysis and interpretation of qualitative data in consumer research. *Journal of Consumer Research*, **21**, 491–503.

Van Maanen, J. (1988) *Tales of the Field: On Writing Ethnography*, University of Chicago Press, Chicago.

Wolcott, H.F. (1999) *Ethnography: A Way of Seeing*, Altamira Press, London.

experimental design

Harmen Oppewal

Experiments focus heavily on *internal validity*, which concerns whether an observed relationship is a true causal relationship (*see* VALIDITY IN EXPERIMENTATION). The aim is to rule out any alternative explanation of the relationship observed in the experiment between the independent variables and the dependent variables (Shadish, Cook, and Campbell, 2002). Suppose we wish to test the effects of a loyalty card on customer spend. Existing company data may suggest that card owners spend more. However, card ownership may be higher among customers with higher income levels. Card ownership is therefore *confounded* with customer income in the company's data and income may well be the real cause for the observed higher spending by card owners.

To rule out such alternative explanations, or "third variables," experimental designs comprise two core features: control and randomization. *Control* means that the researcher actively manipulates the factors of interest while keeping other variables as much as possible constant. To study the effects of a loyalty card we need to keep the card conditions fixed over the study period and retail outlets. It is difficult to fully control all possible variables. This is why experiments are typically conducted in controlled environments. Test conditions are standardized to avoid that other factors can covary and influence the outcomes. Additional variables can be measured to serve as potential statistical controls in the analysis.

Randomization takes over where control cannot be achieved by other means. Test units are assigned to different conditions based on mere chance, for example, by using a probability generator. In our example, we would determine by random draw who receives the loyalty card and who does not. In that case, any observed difference in the outcome between the two conditions can only have been caused by the manipulation, or by the combined effect of all other possible variables on which the test units differed, including, for example, age and gender. Because units were randomly assigned, we can use statistical theory to derive the probability of such differences appearing based on chance alone (*see* CAUSAL RESEARCH). If an observed outcome is too unlikely to be a mere chance effect, we conclude that the only other possible cause is the manipulated difference between the conditions.

The implementation of random assignment requires that the researcher sets the stage and "manipulates" the conditions according to the predetermined plan that includes a procedure for random assignment. This is not always feasible in practice. Applied research often uses existing groups or sites and typically these cannot easily be split and allocated to different conditions. Similarly, it may be difficult, or even unethical to create different conditions (*see* ETHICS IN MARKETING RESEARCH).

Finding the best ways of implementing manipulation and control is the subject of "experimental design." Although, the focus in experiments is on internal validity, which is best achieved under homogeneous and specific circumstances, there is always a trade-off with external validity. Especially in applied research, there is often a strong call for studying "real" situations, and for good reasons. The more similar the test situation to the target situation, the less burden there is for the researcher to show that the results will apply to the target situation. *Field experiments* are experiments that are conducted using real or realistic settings. Sometimes the situation allows maintaining proper levels of control and randomization without any artificiality to the setting. Such studies are however exceptions and depend a lot on the ingenuity of the researchers to find a suitable design.

Often researchers have to live with the design limitations that the study setting or budget imposes and random assignment of individual test units cannot be achieved. This is, for example, the case where existing groups or conditions are utilized, such as when different student classrooms are exposed to different advertisements and asked to rate an advertisement's appeal. When strict randomization cannot be applied, the research design becomes a *"quasi-experimental" design* (Cook and Campbell, 1979). Because no proper randomization can be applied, statistical theory cannot be used to automatically rule out "third

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

variable'' effects. As a consequence, there is greater burden on the researcher to argue explicitly and, where possible, demonstrate that no third variables have influenced the outcomes. (*see* VALIDITY IN EXPERIMENTATION).

Bibliography

Cook, Th.D. and Campbell, D.T. (1979) *Quasi Experimentation: Design and Analysis Issues for Field Settings*, Houghton Mifflin, Boston.

Shadish, W.R., Cook, Th.D., and Campbell, D.T. (2002) *Experimental and Quasi Experimental Designs for Generalized Causal Inference*, Houghton Mifflin, Boston.

purchase, scanner, and media panels

Frank J. Mulhern

Purchase, scanner, and media panels are collections of individuals or households who agree to provide detailed information on their behavior over a period of time. Participants in these panels provide marketers with detailed information on purchase behavior or media use. Panels are designed to represent the demographic and ethnic population of the nation, and range in size from a few thousand to tens of thousands of participants (Sudman and Wansink, 2002).

Purchase Panels

In purchase panels, participants record information about what they buy. Typically, participants, upon returning from a shopping trip, enter their purchase information into either a diary or an online data collection system (*see* WEB SURVEYS). The largest provider of purchase panel information is the NPD Group (www.npd.com). NPD manages an array of panels in over 20 countries. Data from the panels are analyzed and reported by industry type, including automotive, health and beauty, consumer technology, groceries, and so on. Brand marketers obtain information from the panels on the demographic composition of their customers and where they make their purchases.

Purchase panels have the advantage of providing information on essentially everything a household purchases, unlike other data sources which tend to be limited to a few product categories, or a single retailer. Purchase panels allow for analysis and interpretation of what brands people purchase over time, which brands are purchased at which retailers, and how the composition of household spending varies over time.

The disadvantage of purchase panels is that they require the participants to enter information manually, either on paper or a computer. Participants may neglect to enter information for all their purchases for a variety of reasons including carelessness (*see* RESEARCH RELIABILITY AND VALIDITY) and concerns about privacy (*see* INTERNAL DATABASES FOR MARKETING RESEARCH).

Scanner Panels

Scanner panels are a special type of purchase panels, where participants are enrolled in a program that automatically captures data on their purchases at retail stores. Scanner panels overcome the limitation of requiring participants to input information manually. AC Nielsen operates a major scanner panel called *Homescan*, where consumers scan purchases when they return from shopping.

In some cases, the data collection of panel members takes place at stores. Scanner panels can be for a single product category or more broadly for all purchases. Store-based scanner panels capture participants' purchases across several retail chains. Scanner panels tend to be much smaller than purchase panels because the cost of collecting the data is higher since retailer point of sale systems must be integrated into a single process across stores. Scanner panel data is an excellent source of information about repeat buying, brand switching, price sensitivity, and promotion response with multiple regression models (*see* MULTIPLE REGRESSION).

Media Panels

Information on media consumption behavior is necessary for advertisers to successfully place ads where they will be seen by target audiences. That information comes from media panels that consist of households who agree to have their media use tracked by electronic devices. By far, the most important media panel is the Nielsen Television Index (NTI) operated by Nielsen Media Research (www.nielsenmedia.com). The NTI panel consists of 5000 households whose television viewing is captured by an electronic "peoplemeter" that tracks, when the television is on, what channels are watched and who in the household is watching. Data from the NTI measures the size of audiences for TV shows and advertisements and is used to establish prices for advertising placements. Radio consumption is measured by Arbitron (www.arbitron.com) through diary entries about which radio stations participants listen to and in what situations. Print media use is measured not with panels but with data on *circulation*, a measure of distribution, and *readership*, a survey-based (*see* SURVEY RESEARCH) measure

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

of how people say they read newspapers and magazines.

With the growth of the Internet as a medium, services are providing measures of news, entertainment, and advertising consumption online. Nielsen Netratings (www.netratings.com) is one of a variety of services that track and report on Internet use. BIGresearch (www.bigresearch.com) conducts an extensive media panel with over 8000 participants. Participants enter information about media consumption through an online data collection system. BIGresearch provides detailed information on both traditional and online media use in their simultaneous media usage study.

While panels provide excellent information for marketers, they suffer from problems related to sampling (*see* SAMPLING TECHNIQUES), participation, and attrition. While participants are recruited to represent the population at large, panels often have difficulty achieving proper representation. In such cases, the data is statistically adjusted to better represent the population. Attrition, panel members quitting participation, is particularly a problem because it distorts the representativeness of the panel data over time. Panel companies financially compensate participants to minimize attrition.

Bibliography

Sudman, S. and Wansink, B. (2002) *Consumer Panels*, Carnage Publishing.

validity in experimentation

Harmen Oppewal

Types of Validity

Four types of validity are to be considered when evaluating whether research findings can be interpreted as a legitimate test of a hypothesized relationship between an independent and a dependent variable (Cook and Campbell, 1979). They are construct validity, statistical conclusion validity, external validity, and the one at which experiments excel: internal validity. Validity, in general, is discussed in RESEARCH RELIABILITY AND VALIDITY. This article focuses on validity in experimentation.

- **Construct validity** concerns whether the operations and variables in the experiment correctly represent the theoretical constructs as intended. This includes the research setting, the experimental manipulations, and the dependent variable (*see* RESEARCH RELIABILITY AND VALIDITY).
- **Statistical conclusion validity** concerns whether statistical assumptions and procedures are valid and are correctly applied in the analysis. This includes whether randomization has been properly implemented (*see* EXPERIMENTAL DESIGN).
- **External validity** concerns whether the observed relationships can be extrapolated, or generalized, to other samples of units (e.g., people), situations, or times. If a study is conducted in more realistic settings, then there is a smaller gap between the test situation and the setting that actually needs to be understood, making generalizability, which always remains a leap of faith, less of an issue (Shadish, Cook, and Campbell, 2002). An argument that is often used by experimenters is that if no relationship can be established under controlled circumstances there is no point looking at external validity. Once the relationship has been established, further testing is required to assess the "boundary conditions" beyond which the relationship does no longer apply (Calder, Phillips, and Tybout, (1981, 1982); Lynch, 1982).
- **Internal validity** is the extent to which alternative explanations for an observed relationship between variables can be ruled out. In other words, did the independent variable really cause the change in the dependent variable? Through the use of standardized, experimentally controlled conditions, and random assignment of participants to the conditions, experiments allow answering this question with maximum levels of confidence.

Threats to Internal Validity

Although, for any other type of research the internal validity of a relationship has to be explicitly considered and argued for, in properly conducted experiments internal validity is almost automatically guaranteed. Cook and Campbell (1979) identified several types of threats to internal validity.

- **Selection bias** is the most severe internal validity threat. It concerns any systematic difference between the experimental groups not accounted for through standardization or randomization. In experiments, selection is mostly an issue if it occurs in combination with one of the other threats; vice versa the threats below are only a real issue in the experiments if they affect different experimental groups in different ways.
- **History effects** can occur when experimental groups are exposed to different external events during the course of the experiment. The change in the dependent variable can therefore not be uniquely ascribed to the experimental procedure.
- **Maturation effects** are any changes within the experimental units, whether respondents, organizations or sites, that occur during the experiment, and that are not induced by the experimental manipulation. If such changes appear only in particular experimental groups, the experimental results cannot uniquely be attributed to the experimental treatment.
- **Testing effects** occur where exposure to the experimental measures (instead of exposure

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

to the treatment) results in a change in the experimental units. If respondents in an ad test, before seeing the ad, are asked about their brand perceptions, the mere exposure to these perception questions may result in improved brand recall. This improvement should not be attributed to the ad as it is an effect of having been asked questions about the brand.

- **Instrument variation effects** may occur when different measurement instruments are used for different experimental groups resulting in differences in the dependent variable, for example, if in one condition respondents answer a paper-and-pencil task while in another they answer an online task.
- **Statistical regression** occurs if participants are selected for different conditions on the basis of their performance on some test. Because of the measurement error in the pretest, there will be a natural tendency for participants with high scores to display somewhat lower scores if a similar test is conducted again, for example, after the experimental manipulation. The opposite will happen to participants with a low pretest score.
- **Experimental mortality** concerns whether respondents complete the experimental tasks. If in a satisfaction study one condition takes longer to complete this may result in fewer people completing this condition. If these dropouts are the ones that are less satisfied with the product, the result will be that the remaining group achieves higher scores than they would have otherwise. Clearly, this is an artifact and not evidence that satisfaction has increased (*see* CUSTOMER-SATISFACTION RESEARCH).

Bibliography

Calder, B.J., Phillips, L.W., and Tybout, A.M. (1981) Designing research for application. *Journal of Consumer Research*, **8**, 197–207.

Calder, B.J., Phillips, L.W., and Tybout, A.M. (1982) The concept of external validity. *Journal of Consumer Research*, **9**, 240–244.

Cook, Th.D. and Campbell, D.T. (1979) *Quasi Experimentation: Design and Analysis Issues for Field Settings*, Houghton Mifflin, Boston.

Lynch, J.G. Jr. (1982) On the external validity of experiments in consumer research. *Journal of Consumer Research*, **9**, 225–239.

Shadish, W.R., Cook, Th.D., and Campbell, D.T. (2002) *Experimental and Quasi Experimental Designs for Generalized Causal Inference*, Houghton Mifflin, Boston.

web surveys

Tuo Wang

A web survey is a self-administered survey conducted on the Internet or through a website. Instead of sending a questionnaire to a respondent, the questionnaire is stored electronically on a web server. To complete the questionnaire, a respondent, usually clicks a link that leads to the web server. When the respondent finishes the survey, the survey data is sent to a server via the Internet. Researchers can retrieve the stored survey data from the server anytime via the Internet. Web surveys can be conducted either through invitation by the researcher (often delivered via e-mail) or by a "pop-up" window that appears when a potential respondent visits a certain website. The percentage of US households with Internet access has been increasing rapidly. Web surveys have become more attractive to marketing researchers owing to the increasing representativeness of web survey respondents. If the sample for a web survey is carefully selected and fits the target population, a web survey can provide both efficiency and quality. Web surveys enjoy unbeatable advantages over traditional (snail) mail surveys in terms of saving time (e.g., mailing and data entry) and money (e.g., costs associated with printing, mailing, and data entry). When interactive technologies are properly applied, a web survey can also enhance questionnaire flexibility and provide visual stimuli to the respondents. The limitations of the web survey are (i) the requirement of easy access to the Internet for respondents and potential sampling frame bias (when the Internet is not easily available to certain segments of the target population); and (ii) the unknown quality of a web survey's results when an interest group can potentially manipulate responses or when Internet security may be compromised. Another concern when using web surveys is the potentiality of low response rate when compared to a mail survey Cook, Heath, and Thompson, (2000). This may be a serious problem for some web surveys when the soliciting emails are classified as "spam" or "junk mail". Nonresponse bias is a problem that web surveys need to overcome, either through "propensity scoring," which gives more weight to underrepresented respondents, or through the use of a complementary data collection method, such as a traditional mail survey.

To conduct a web survey, researchers can either design the survey using web-based programming language (e.g., HTML) or using commercial software that enables them to create a web-based survey without programming skills (e.g., surveymonkey.com). In summary, web surveys are gaining popularity as the Internet becomes easily available to a larger population. Web surveys save time and money for the researcher. According to Research Industry Trend 2006, web-based surveys are now the most popular survey method with 36.8% of total usage.

Bibliography

Cook, C., Heath, F. and Thompson, R.L. (2000) A meta-analysis of response rates in web- or Internet-based surveys. *Educational and Psychological Measurement*, **60** (6), 821–836.

Research Industry Trends, (2006) Report, Pioneer Marketing Research, GreenBook, Dialtek and Rockhopper Research.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

comparative evaluation of survey methods

Tuo Wang

There are generally four criteria used to evaluate a survey method: speed of data collection, cost, data quality, and flexibility. The choice of an appropriate survey method thus depends on answers to the following questions: (i) How soon does the researcher need the information? (ii) How much can the researcher spend on data collection? (iii) What is the data quality desired by the researcher? (iv) How much flexibility is desired by the researcher?

How Soon does the Researcher Need the Data?

When time is short, the researcher may have to compromise quality for speed. Time to complete a survey includes deployment time, respondent recruitment time, and response time. Response time is likely to be the most unpredictable when a survey is self-administered. All things considered, the choice of a survey method when speed is a priority can be evaluated as follows:

Fast methods.

1. *Mall-intercept interviews*: This also includes intercepting shoppers in other high-traffic locations. Owing to the high density of shoppers in these locations, as well as the high response rate for personal interviews, this method can potentially generate a high number of survey respondents in a short time period. When the demographic profile of the target population needed for the research is similar to a mall's shoppers, this approach can save both time and money for the researcher without significantly compromising data quality.
2. *Web surveys*: These can be very fast if respondents are provided with the right incentives to complete the survey. There is no delay between sending and receiving the data.
3. *Telephone interviews*: Without knowing the availability of a potential respondent (call blocking, bad timing, etc.), it may take additional time to recruit qualified respondents with the first filter question. Once a respondent is contacted and agrees to participate, researchers can quickly get the response without further delay.

Slow methods.

1. *Personal interviews*: These require extra travel time for the interviewer.
2. *Mail surveys*: This is the slowest survey method, primarily owing to the long lead time for preparation and deployment, as well as potential response delays by subjects. A reminder or incentive needs to be provided to reduce response delay.

It is important to recognize that there are a large number of vendors that provide specialized services to collect data by using different survey methods. Major vendors of research services are listed in "The Green Book". The response time, or time to collect necessary data, can be reduced by hiring outside vendors. For example, the response time for an online survey may range from one day to two weeks. For a mail survey, it may take an average of one month to obtain data from those who are likely to respond. An outside vendor can easily reduce the response time by hiring an adequate number of interviewers to collect data through face-to-face interviews or telephone interviews. Vendors can collect data within weeks from these methods. Thus, response time of a survey method is dependent upon the budget available for collecting the data.

How Much Can the Researcher Spend on Data Collection?

Generally speaking, the least expensive method per interview is the web survey (no interviewers cost, no mailing cost, and no data entry cost) McDaniel and Gates, 2007, p. 181. Mail surveys are traditionally economical choices in the absence of interviewers. Telephone interviews are also reasonable compared to personal interviews, but are losing ground to web surveys owing to the availability of Internet access in more households. The most expensive is personal interviews, which involve both travel cost and compensation for quality interviewers.

When evaluating costs of using different methods, one needs to separate and consider the fixed and the variable cost per survey. The major fixed cost involves training/hiring interviewers, buying/renting software to design

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

the survey, and hiring skilled/unskilled labor. The variable costs involve the cost of buying mailing or e-mail addresses, postal charges, and telephone numbers (Groves, 2004). In general, except for the face-to-face interview, the variable cost per completed survey does not exceed a few dollars per survey, whereas the telephone and face-to-face interview result in high fixed costs owing to the need for skilled personnel to conduct the interviews.

What is the Desired Data Quality?

Any survey method is prone to error. The researcher tries to minimize the systematic error (or bias). If a survey is conducted by an interviewer, interviewer bias is likely to be present. If a survey is self-administrated, nonresponse bias may be serious and needs to be addressed (Groves, 2004). A general rule of thumb is (i) a more expensive survey method generates better data (better interviewers, better sampling, etc.); (ii) a survey that is more time consuming generates better data (respondents have more time to think about the questions or to interact with interviewers).

How much Flexibility is Needed by the Researcher?

Again, personal interviews are the choice when maximum flexibility is required (Aaker, Kumar, and Day 2001, p. 237). Personal attention from a good interviewer ensures flexibility even when the researcher is not aware of, or cannot anticipate, the potential problems before the interview/survey is conducted. When differences among respondents can be classified before the survey is administered, web surveys are a good choice, having the ability to deliver a high level of personalization (the survey questions can be customized easily on the basis of respondents' answers to previous questions). Mail surveys or web surveys are better choices than telephone interviews when pictures are required. Mall-intercept interviews are ideal when physical products need to be demonstrated or sampled.

Please see SURVEY RESEARCH for further details regarding different survey methods.

Bibliography

Aaker, D.A., Kumar, V., and Day, G.S. (2001) *Marketing Research*, 7th edn., Wiley.

Groves, R.M. (2004) *Survey Errors and Survey Costs*, Wiley.

McDaniel, C. and Gates, R. (2007) *Marketing Research*, 7th edn., Wiley.

comparative scaling technique

Vanitha Swaminathan

Scaling is the process by which questions are used to assign numerical values to measure subjective properties of an object. In marketing, scales are used to quantify a variety of aspects, including brand attitudes and brand preferences, liking for various ads, and so on. Sometimes, scales can be used to measure the extent to which two or more entities can be compared or evaluated, for example, advertisements or brands.

There are four types of comparative scales: (i) paired comparison; (ii) rank order; (iii) constant sum; and (iv) *Q*-sort.

1. *Paired Comparison*: A paired comparison is a specific type of comparative scaling technique. In a paired comparison, two versions of a new product may be tested against each other. Respondents can be surveyed following a product use involving the two versions of the new product, as described below.
   Which of these two brands – Breyer's or Ben & Jerry's do you think is better?
   ☐ Breyer's ☐ Ben & Jerry's
2. *Rank order*: In this type of comparative scaling method, respondents are asked to rank order various options as shown below. Rank each of the following brand names from most preferred to least preferred (1 = most preferred and 5 = least preferred):

	Ranking
Breyer's	_____
Ben & Jerry's	_____
Häagen-Dazs	_____
Klondike	_____
Turkey Hill	_____

3. *Constant sum*: In the constant sum scale, respondents are asked to assign points to each of the options with more points assigned to the more preferred option.
   Please assign 100 points across each of the following brands based on your relative preference for each of these brands (please note that the total should sum to 100).

	Assign Points (0–100)
Breyer's	
Ben & Jerry's	
Häagen-Dazs	
Klondike	
Turkey Hill	
Total	**100**

4. *Q-Sort*: This is a rank order procedure in which respondents are asked to sort a given number of items or statements and classify them into a predetermined number of sets (usually 11) according to some criterion such as preference, attitude, or behavioral intent. Typically, respondents are given cards to sort. This is typically used when there are a large number of items that should be ranked within a short period of time.

The advantage of comparative scaling is that it forces respondents to compare the options and focus attention on the relative merits and demerits of the options being compared. This is particularly useful in situations in which the marketer desires to identify competitive differences between various products and ads. Comparative scales are also particularly beneficial in developing countries with lower education levels and sophistication levels.

One disadvantage is that when the entities being compared are either vastly different or when there is no basis for comparing the entities (e.g., a fruit juice versus a carbonated soda) – the results of this may be inconclusive. Another disadvantage is that as the number of entities (e.g., brands/products) increases, the number of comparisons that the respondent must make also increases. For instance, if there are four brands that must be compared, a total of $(4 \times 3)/2 = 6$ comparisons should be made. This may cause respondents to become bored or tired, which can reduce their ability to discriminate between options.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

When comparing values on different variables, it is often assumed that differences in scores can be used as an indicator of relative preference. However, there are biases, for example, acquiescence bias or extreme response style bias, which may affect the answers (*see also* SOCIAL DESIRABILITY BIAS; COMMON METHODS BIAS). Acquiescence bias is the tendency for a respondent to agree with a response. A set response is the tendency for a respondent to answer a series of questions on a certain direction regardless of their content. Response biases may result in responses that are similar across all the options that are compared. Comparative scaling methods may reduce these biases by forcing respondents to make explicit comparisons across various options.

Bibliography

Churchill, G.A. Jr., and Tom, J.B. (2007) *Basic Marketing Research*, 6th edn, Thomson-Southwestern College Publishing, Mason, OH.

Zikmund, W.J., and Barry, J.B. (2007) *Exploring Marketing Research*, 9th edn, Thomson-Southwestern College Publishing, Mason, OH.

electronic scanner services

Frank J. Mulhern

Electronic scanner services capture brand sales data from scanner systems at retail store checkouts and sell that information to manufacturers and retailers. The retail scanner systems read the universal product codes (barcodes) on each item purchased and tabulate information on brand sales. Unlike scanner panels (*see* PURCHASE, SCANNER, AND MEDIA PANELS), which capture data on the purchases of individuals and households, electronic scanner services provide aggregate data for individual store locations, retail chains, and metropolitan areas. The scanner systems provide data with a great deal of detail including brand names, package sizes, and prices.

Two companies dominate the electronic scanner services industry – AC Nielsen (www.acnielsen.com) and Information Resources Inc. (www.infores.com). Both companies collect data from carefully selected samples of retail store locations and conduct statistical analysis to estimate brand sales at the account (retail chain) and market levels. While the primary emphasis is on grocery stores, both companies also obtain sales data from major drug and convenience stores. Data is collected on an ongoing basis and aggregated into weeks to match the retail practice of setting price and promotions on a weekly basis. In addition to selling data, the service providers provide some analytic services such as estimating a brand's weekly baseline (nonpromotion) sales and incremental sales due to price discounting.

Nielsen's SCANTRACK service collects weekly data from over 4800 stores representing over 800 retail chains in 52 market areas in the United States. The most important information provided by these services is the timely tracking of brand sales and brand market share at the point of purchase. The information, which first became available in the 1980s after stores installed scanner checkout systems, has revolutionized marketing mix planning by enabling manufacturers to have precise information on how brand sales and market share are affected by prices, promotions, and advertising.

The availability of scanner data has had an enormous impact on point-of-purchase marketing, which includes shelf-space allocations and in-store advertising and promotion. Manufacturers use sophisticated statistical procedures to assess the effect of a variety of marketing mix elements including prices, in-store coupons, end-of-aisle displays, retail cooperative print advertising, and digital signs in stores. The statistical analyses reveal the impact of each marketing variable on brand sales and market share, allowing brand managers to better develop marketing practices. Statistical analysis of scanner data also reveals patterns of brand switching, variety seeking, price sensitivity, and promotion response.

Because the electronic scanner services provide information on sales in major cities and other geographic areas, manufacturers can use the data to evaluate the effects of local and regional media advertising. In fact, many manufacturers conduct field experiments (*see* EXPERIMENTAL DESIGN) where they vary advertising or promotion in different geographic areas, then evaluate scanner data on brand sales in those areas to assess the effect of the marketing.

While scanner data is of great value for marketers, it has some significant limitations. The quality of scanner data is quite good, but occasional data-collection errors occur when store clerks are unable to scan a barcode or a data-input error assigns the wrong price to an item. Since scanner data aggregates brand sales in a store, it contains no information on the identity of customers or the basket of products purchased together. Also, there is no information on household purchasing across locations of a single chain or across retail chains. These disadvantages are diminishing as the electronic scanner services move to household-level data made available by grocery store buyer clubs. The premier provider of information from that data is DunnHumby, Inc (www.dunnhumby.com), which pioneered the analysis of household-level grocery purchase data at Tesco in the United Kingdom (Humby, Hunt, and Phillips, 2003).

Bibliography

Humby, C., Hunt, T., and Phillips, T. (2003) *Scoring Points: How Tesco Continues to Win with Customer Loyalty*. Kogan Pae Ltd, London.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

field work/data collection process

Pamela Grimm

Field work refers to the collection of primary data from consumers or business-to-business customers through survey or observation. Unlike the experimental studies that are generally conducted in a controlled environment, researchers studying marketing often rely on data collected in "the field," – outside a laboratory and in a more natural setting, such as face-to-face, telephone, or online surveys conducted at the respondent's residence or in shopping malls. Data may also be collected through observations of the respondent's behavior in the home or in some other setting such as a retail outlet. Therefore, for the purpose of this section, *field work* is defined as the process of data collection using any one of several survey methods such as face-to-face interviews; telephone, postal, and online surveys; or observation.

Most firms subcontract with other firms for data collection. There are a number of vendors who specialize in data collection, many of them listed in a publication known as *The Green Book*. This is a resource compiled by the American Marketing Association that provides a directory of research vendors. Although contracting the whole research study or subcontracting a part of it, such as data collection, relieves the researcher from the headache of doing the research study or collecting data, not every firm can afford to contract or subcontract the entire, or part of, the research study. It is also desirable for a researcher to have sufficient knowledge of the research and data collection processes to be able to interact intelligently with a subcontractor. As a result, a researcher needs to have a sense of what collecting data for a research study entails.

Once the survey instruments are pretested and revised, the researcher is ready to collect the actual data to test hypotheses or answer research questions. If the survey is conducted by mail using the postal service, the process of mailing out surveys primarily involves manual labor rather than specialized training. Of course, people who are mailing the surveys need to ensure that each respondent receives the cover letter, questionnaire, incentives (if any), and means of returning the completed survey without incurring any cost (e.g., prestamped envelope).

Online surveys can be "fielded" (i.e., launched) even more quickly than snail mail surveys. Online data collection is typically done through software controlled by a service provider, such as Qualtrics, Survey Monkey, or Zoomerang. All that is needed is a list of valid respondent e-mails for the target population and an online survey.

A field study that is being conducted through an interviewer, either face to face or over the telephone, will require training of the interviewers. Training that might be required could include sampling of respondents, review and practice with the survey instrument, recording of responses, special instructions (do's and don'ts), and so on. Training of interviewers in a field study is important because of the potential for contamination of data due to a failure to follow the correct data collection protocol. Of course, interviewers could also introduce errors in data owing to both verbal (the way they ask questions) and nonverbal (body language, gender) indicators.

A special concern when conducting a field study must be security. Because this research is typically conducted outside the researchers premises, it is very important that steps be taken to ensure the safety of completed surveys or data files. This is important to protect the researcher's proprietary information and the respondent's privacy and any guarantee given of confidentiality of responses.

Bibliography

De Leeuw, E. D., Hox, J. J. and Dillman, D. A. (2008) *International handbook of Survey Methodology*, European Association of Methodology, Lawrence Erlbaum Associates, New York.

New York AMA communications Services, Inc. (2009) *GreenBook, Volume I and II – Worldwide Directory of Marketing Research Companies and Services*, New York.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

noncomparative scaling technique

Vanitha Swaminathan

Noncomparative scaling techniques is a set of scaling methods in which each item is scaled independently of each of the others. There are various types of noncomparative scaling techniques such as the graphic rating scale, Likert scale, semantic differential scale, and Stapel scale (*see* ITEMIZED RATING SCALES (LIKERT, SEMANTIC DIFFERENTIAL, AND STAPEL)).

A graphic rating scale involves asking respondents to place a check mark on a continuous line. Only the endpoints of a graphic rating scale are numbered. This type of a continuous rating scale enables respondents to provide a wide range of responses. For example,

Listed in Annexure 1 is a set of attributes. Please evaluate how important each attribute is to you in your decision to purchase of a car. Place an X on the horizontal line to indicate how important each attribute is to you.

The Likert, semantic differential, and Stapel scales are examples of noncomparative-itemized rating scales (*see* ITEMIZED RATING SCALES (LIKERT, SEMANTIC DIFFERENTIAL, AND STAPEL)). In these types of scales, there are discrete points along the scale that are typically either numbered or have verbal labels associated with each of the discrete points.

In a Likert scale (also known as an *agree–disagree scale*), respondents are asked to indicate their level of agreement from strongly agree–strongly disagree on either a five- or a seven-point scale. An example of the Likert scale is given in Annexure 2.

An advantage with the Likert scale (as shown above) is that various scale items can be combined. In the above example, respondents' answers to each of the statements can be combined to form a summated score toward Amazon.com.

The semantic differential scale, originally developed by Osgood, Suci, and Tannenbaum (1957), involves providing respondents with five- or seven-point scales with bipolar adjectives at either end. An example of a semantic differential scale is provided below:

Please rate Southwest Airlines on the dimensions as per Annexure 3. (Place an "X" suitably on each).

Semantic differential scales are used in marketing primarily to measure brand or store image. The results from several semantic differential scale items can be combined to create a visual image profile of store or brand image for a given store or brand. An example of such a visual image profile is given below:

Comparing the brand image of Southwest airlines and American Airlines. See Annexure 4.

The advantage of the semantic differential scale is that it allows marketing researchers to visually depict the scores across a variety of entities (stores, brands, and companies). One disadvantage may be that respondents have been shown to favor the middle points of a scale. As a result, extreme positive or negative ends of the scale are not used by respondents. Another disadvantage could be that researchers assume that data derived from semantic differential scales is interval data, although, the above scale could sometimes only have ordinal properties.

The Stapel scale helps measure direction as well as intensity of response. A Stapel scale involves placing a single adjective in the center of an even number of values (e.g., ranging from −5 to +5). Respondents are asked to indicate how close to or distant from the adjective a given stimulus may be. The advantage of the Stapel scale is that it is a simple scale which can easily be incorporated in telephone interviews. Compared to the semantic differential scale, it is easier to construct. An example of a Stapel scale

Annexure 1

Gas mileage	Not important	-----	Very important
Brand name	Not important	-----	Very important
Price	Not important	-----	Very important
Reliability	Not important	-----	Very important

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Annexure 2

	Strongly Disagree	*Disagree*	*Neither Agree Nor Disagree*	*Agree*	*Strongly Agree*
Amazon.com is very responsive to customers	——	——	——	——	——
Amazon.com maintains privacy of customer information	——	——	——	——	——
Amazon.com makes it easy to return merchandise	——	——	——	——	——
Amazon.com has a website that is easy to navigate	——	——	——	——	——

Annexure 3

Low value for money	— : — : — : — : — : — : —	High value for money
Typically late	— : — : — : — : — : — : —	Typically on time
Poor baggage handling	— : — : — : — : — : — : —	Excellent baggage handling
Bad in-flight service	— : — : — : — : — : — : —	Good in-flight service
Poorly maintained planes	— : — : — : — : — : — : —	Well-maintained planes

is given below:

+5
+4
+3
+2
+1
Wide product assortment
−1
−2
−3
−4
−5

Please indicate the extent to which Target has a wide product assortment.

If you believe the adjective describes the store accurately, then choose a higher plus number.

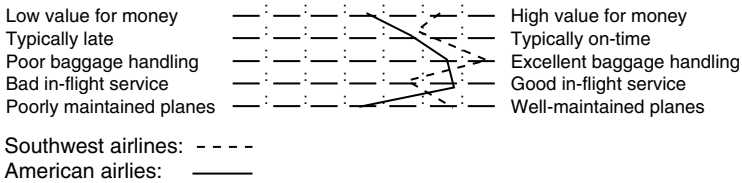
If you believe that the adjective does not describe the store, then choose a lower minus number.

In addition to deciding on the type of scale (on the basis of the above description), the researcher must also decide on the following:

1. *Number of scale categories*: The larger the number of scale response categories, the more the information. However, having too many categories makes it harder for the respondents to react. Typically, market-research surveys (*see* SURVEY RESEARCH) measure brand attitudes or brand preferences using five-, seven- or nine-point scale categories.
2. *Odd/Even number of response categories*: A key decision that a researcher has to make is whether to have odd/even number of response categories. Choosing an odd number of response categories can help respondents by providing them a midpoint, which can help capture "neutral" responses. An even number of response categories forces respondents to provide a directional response, and may sometimes artificially force neutral respondents to skew their responses to suit the scale.
3. *Balanced versus unbalanced scales*: A key decision is whether the questions have to have equal number of favorable or unfavorable responses. In unbalanced scales, there are an unequal number of favorable and unfavorable response categories.
4. *Forced versus nonforced scales*: In forced scales, the respondents cannot have the option of expressing either "no opinion"

Annexure 4

Comparing the brand image of Southwest airlines and American airlines



or "do not know" in response to a question. Forced response scales may be problematic in cases where a large number of respondents have insufficient knowledge to respond.

5. *Nature and degree of verbal descriptors*: Verbal descriptors are particularly helpful to respondents since they enable them to articulate the responses better. However, researchers must exercise care in ensuring that verbal descriptors are unambiguous and accurately reflect the response categories.

Bibliography

Churchill, G.A. Jr., and Tom, J.B. (2007) *Basic Marketing Research*, 6th edn, Thomson-Southwestern College Publishing, Mason, OH. (semantic differential scale adapted from chapter 13, p. 285).

Osgood, C.E., Suci, G.J., and Tannenbaum, P.H. (1957) *The measurement of meaning*, The University of Illinois Press, Urbana, IL.

Zikmund, W.J., and Barry, J.B. (2007) *Exploring Marketing Research*, 9th edn, Thomson-Southwestern College Publishing, Mason, OH.

nonprobability sampling

Richard J. Fox

Nonprobability sampling entails judgmental selection of a subset from a population in such a way that the subset selected is not dictated by chance or randomness. Sampling is performed to learn about a population, but when nonprobability sampling is used, generalizing the results to the entire population is questionable. Further, the statistical methods related to measuring the precision and accuracy of estimates derived from the sample are no longer valid. However, nonprobability samples can be useful for gaining general understanding about how a population behaves, how opinion on a topic varies from individual to individual, and so on.

There are a number of types of nonprobability sampling (see Malhotra (2007) for a discussion of nonprobability sampling techniques). As the name suggests, a convenience sample is one that is convenient to the individuals selecting the sample. For example, a college student might question his or her friends about a movie that they have seen in order to determine whether he or she might be interested in seeing it. A company might enquire from its major clients, namely, those with whom the company does considerable business, what they think of a new billing procedure. Student samples, often used by university professors in academic research, are also convenience samples.

Judgment sampling is another form of nonprobability sampling. In this case, the investigator or individuals seeking the information about a population subjectively choose members from the population to compose the sample. For example, business managers might decide to interview large customers regarding their plans for the next fiscal year. The rationale is that information from large customers is critical to planning for the coming year. Product engineers in a software company may choose to talk to the "lead users" of the various software tools that the company markets. The reasoning is that future development will likely be driven by these lead users.

Snowball sampling, another form of nonprobability sampling, involves asking individuals initially selected to be in a sample to provide contact information for others to be potentially included in the sample. This is a particularly useful tool when looking for individuals whose incidence is very low. For example, to obtain a sample of individuals with extensive hiking experience, a person who qualifies could be asked to provide contact information regarding others who have similar interest or expertise.

A self-selecting sample is also a form of nonprobability sample. In a self-selecting sample, individuals "volunteer" to be in the sample. For example, cell phone text polls provide a vehicle for individuals to "text" a vote for one of two or more options. The votes are tallied and the winner announced during the event or broadcast in which the poll is conducted. The only people included in such a sample are those who choose to participate. These are likely to be individuals who have a vested interest in the topic, and hence are not necessarily a representative sample of the population. Even if the number of responses obtained is extremely large, the validity of the sample is still in question. In a large study of adults in the United States, the majority of people reported that polls of this type are valid because many people participate. Again, this is a misconception. The sample size may be large but the sample may be far from representative.

Finally, another form of nonprobability sampling is quota sampling. In this instance, restrictions are placed on the sample so that the numbers of sample units classified into specific categories meet prescribed requirements. For example, the sample may be required to be half male and half female, or to match the US age distribution. However, there is no assurance that the sampling method used to meet each quota requirement is such that representative samples are selected for the quota groups. Convenience samples might be used to fill the quotas. Hence, estimates relating to individual quota groups as well as overall estimates are not valid, and their accuracy and precision cannot be determined.

Bibliography

Malhotra, N.K. (2007) *Marketing Research : An Applied Orientation*, 5th edn., Pearson Prentice Hall, Upper Saddle River, NJ.

Churchill, G.A., Jr. and Brown, T.J. (2004) *Basic Marketing Research*, 5th edn., Thompson South-Western, Manson, Oh.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

personal observation

Linda L. Price

Observational approaches have a long history in social science including marketing and consumer research. Broadly speaking, observation methods include all data collection in which participants are not asked to give us data, but rather are monitored for data. Observation can be quantitative (such as head counts or measurements of wear or use) or qualitative (such as participant observation or netnography) (*see* OBSERVATION METHODS). *Personal observation* is a subset of observational approaches in which the researcher is the primary instrument for monitoring participants. For example, studies that use technologies to measure participants' physiological responses to stimuli are observational, but not personal observation. Nonetheless, reliance on technology for personal observation is often substantial. For example, videotaping, cameras, websites, and so on are commonly employed in fieldwork by ethnographers and netnographers. What distinguishes personal observation from other observational methods is that the researcher views the videos, examines the pictures, and visits the community websites in order to collect relevant data and then interprets and represents it in study findings.

Personal observation can be covert or overt, participatory or nonparticipatory and varies in how obtrusive it is in participants' lives. For example, a researcher may seek entry into an activity or community with the explicit aim of doing research (overt), may participate in the practices of that community (participatory), and as a consequence, may change that community in some ways (obtrusive). Alternatively, the researcher may place cameras at a location in a store (overt), observe customer behaviors at that site from a distance (nonparticipatory) and leave participants' unaware of researcher presence (nonobtrusive). In practice, researchers would often employ a mix of participatory and nonparticipatory observations during their fieldwork (*see* FIELD WORK/DATA COLLECTION PROCESS). Personal observation can be more or less systematic in terms of sampling and recording, but should always involve detailed field notes or tallies recorded as close to the moment of observation as possible. By immediately recording observations researchers avoid memory biases that could lead them to selectively remember and forget, generally in ways that systematically conform to their cultural expectations.

Although personal observation can be quantitative or qualitative, in marketing and consumer research it is most often associated with ethnographic research, a qualitative technique. In fact, participant observation is generally considered the central and defining feature of ETHNOGRAPIC RESEARCH. *Participant observation* is a systemic recording and analysis of information gained while engaged in and observing activities, rituals, interactions, and events in order to uncover explicit and tacit aspects of the field of study. The approach that ethnographers use in participant observation is highly individualistic affected by a complex mix of context factors, theoretical frameworks, and personal characteristics. Participant observation is both a data collection and an analytic tool, since participating in activities and events in the field provides researchers with a tacit understanding that shapes both observations and subsequent interpretation of data collected. Multivoiced participant observation that uses multiple researchers in the field participating and observing the same or similar activities, practices, and events can uncover a fuller array of perspectives and often provide a more integrative and complete narrative account. For example, multigendered teams can help give voice to both female and male lived experiences. Member checks and the use of multiple types and sources of data can also increase the perceived trustworthiness of the researcher account. A number of different criteria might be used to evaluate the trustworthiness of the observational account including adequacy, resonance, integration, innovation, and usefulness (*see* RESEARCH RELIABILITY AND VALIDITY).

Personal observation raises many important ethical issues for researchers. These issues include establishing limits to participation, finding the proper mix of observation and participation, and maintaining the anonymity of informants' studied. As with other social science research, a good starting point is to work to

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

ensure that people studied are not harmed by researcher involvements or negatively affected by the information collected and written about them (*see* ETHICS IN MARKETING RESEARCH). Some researchers adopt a critical ethnography or action research perspective and advocate the explicit goal of improving human well-being through participatory research (*see* CONSUMER WELL-BEING). For example, they might argue that the studied community should be the source of research problems and the primary beneficiary of the research.

Bibliography

Atkinson, P. and Hammersley, M. (2006) *Ethnography: Principles in Practice*, 3rd edn, Routledge Press, New York.

Dewalt, K.M., Dewalt, B.R., and Wayland, C.B. (1998) Participant observation, in *Handbook of Methods in Cultural Anthropology* (H.R. Bernard), Altamira Press, Walnut Creek, pp. 259–299.

Emerson, R.M., Fretz, R.I., and Shaw, L.L. (1995) *Writing Ethnographic Fieldnotes*, University of Chicago Press, Chicago.

Johnson, A. and Sackett, R. (1998) Direct systematic observation of behavior, in *Handbook of Methods in Cultural Anthropology* (H.R. Bernard), Altamira Press, Walnut Creek, pp. 301–331.

Kozinets, R.V. (2002) The field behind the screen: using netnography for marketing research in online communities. *Journal of Marketing Research*, **39** (1), 61–72.

Ozanne, J.L. and Saatcioglu, B. (2008) Participatory action research. *Journal of Consumer Research*, **35**, 423–439.

Peñaloza, L. (1994) Atravesando fronteras/border crossings: a critical ethnographic study of the consumer acculturation of mexican immigrants. *Journal of Consumer Research*, **21**, 32–53.

Spiggle, S. (1994) Analysis and interpretation of qualitative data in consumer research. *Journal of Consumer Research*, **21**, 491–503.

Pretesting a questionnaire

Pamela Grimm

Pretesting or pilot testing of a survey is conducted to test the survey instrument and data collection procedure before data collection begins. The objective is to ensure that the questions being asked accurately reflect the information the researcher desires and that the respondent can and will answer the questions. In general, the types of problems that may be encountered in a survey range from misinterpretation of questions to the inability or unwillingness to answer questions or to the length of time required to respond. The number, wording, and layout or ordering of questions may be the root cause of many survey problems.

Pretesting should be considered a critical step to go through when collecting data via a survey. Failure to pretest could increase both the sampling errors (because of nonresponse to either a few questions or the whole questionnaire) and nonsampling errors (because of misunderstanding some questions, incorrect or unclear skip patterns, or any number of other reasons). Pretesting of the questionnaire should replicate, as much as possible, the way the final study will be conducted. Pretesting is usually done on a small sample (about 30) that matches the characteristics of the population that will be sampled for the study. However, less formal pretesting can be conducted with a smaller convenience sample.

Typically, the researcher will ask the respondent to complete the survey and then ask the respondent questions, either one-on-one or in small groups, during a debriefing session. If the survey was administered by an interviewer, as would be the case for telephone or mall intercept surveys, interviewers also are debriefed. An alternative to debriefing after administering the questionnaire is to have respondents "think out loud" while filling out the questionnaire. Both mechanisms provide the researcher with a respondent's reaction to the questionnaire.

While the emphasis is often on qualitative feedback during a debriefing session, it is also important to tabulate the data collected though the pilot test and examine the patterns of responses to each question, looking for questions that respondents are refusing to answer, incorrectly skipping, or which seem to be misinterpreted. These data, however, are not included with the data collected later in the final survey.

More specifically, the researcher addresses the following items in a pretest:

1. Did the respondents clearly understand what was being asked and the terminology being used?
2. Was there any major option missing in any of the questions? If several respondents write new option(s) to a question, options are obviously not exhaustive.
3. Was the respondent willing to answer each question? Were any questions offensive, sensitive, or too difficult?
4. Did the respondent understand the instructions for responding to each question? For example, instructions may not be clear regarding choice of only one option or multiple options to a question.
5. Did the respondent feel that the researcher wanted him or her to answer in some specific way (i.e., did the respondent feel some of the questions were biased)?
6. Did the respondents understand any instructions regarding skipping questions depending upon the answer to previous questions?
7. Did the flow and language of the questionnaire seem logical and natural?
8. Did the time taken to answer the questionnaire seem reasonable?
9. If the survey was administered by an interviewer, did the respondent feel that the interviewer, either verbally or nonverbally, tried to bias his or her response?

This is a short list of areas that the researcher might check during a pretest. Once the pretest is complete, any changes necessary should be made. If changes are substantive, the researcher may want to engage in a second round of pretesting. More details on pretesting questionnaire can be found in QUESTIONNAIRE DESIGN.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

primary scales of measurement

Frederick Winter

Scales in marketing are similar to the trusty ruler. They are used to measure things. But unlike a ruler, which is much like another ruler, different scales are used by marketers to measure different things. It is crucial when using market research to look at data and understand the scale as well as some properties that led to the data that one is viewing.

Scales can be of the following types: nominal, ordinal, interval, and ratio based. The order described also shows the increasing data content as one goes from nominal up to ratio. At the low end of the scale, is a nominal scale that comes from the Latin word *nomeni*, which means "name." The most common example of a nominal scale is the number on a football jersey. Marketers use nominal scale in a number of ways. For example, we could describe the consumer's favorite brand of cereal, and that would represent nominal data. We could indicate the consumer's preferred store where he/she would buy a product. That too would be nominal. There are a lot of statistics that we are used to using such as the average, which are no longer meaningful when we have nominal data. Let us say five consumers bought Cheerios, three consumers bought Frosted Flakes, and two consumers bought Grape-Nuts. The average brand that the group bought makes no sense. Rather some statistic that indicates the norm is the *mode* – which is the most frequently occurring value. In our cereal example, we would say the mode is Cheerios since it is the most frequently occurring value. Nominal data can also be analyzed in terms of the relationship between two or more nominal variables. For example, one might want to know if the preferred breakfast cereal is related to gender, another nominal variable. A common method used by marketing researchers is CROSS-TABULATION which reports joint frequency; association can be tested using the chi-square test statistic to indicate if there is such a relationship between two nominal variables.

If we have asked the consumer to rank order his/her preference for stores we would have an ordinal scale. Ordinal scales contain more content than do nominal scales because it says that something ranked 1 is higher (or lower) in attribute than something that is ranked number 2 and 2 is more (or less) than something ranked 3. Ordinal scale, as the name implies, indicates something about order. (as indicated 1 could be high – such as weight, or low – such as lightness). With ordinal data, there is no implication of the distance between the data. Like a running race, the difference between first and second place might not be the same as the difference between 40th and 41st place. Theoretically, we can report the mode on ordinal data, assuming it is not place in a race (where everybody has a unique value and all places qualify as the mode). Quite often, we summarize the central tendency with ordinal data in terms of the median in which half of the numbers fall above and half below. We can also associate two or more ordinal variables using cross classification, but the techniques available are a little more powerful because the data are more powerful and we can use things like the Spearman rank order correlation coefficient to indicate a relationship between two or more variables.

The data content increases considerably as we go from ordinal data to interval data. In interval data the difference between 1 and 2, which is 1, is the same as the difference between 41 and 42. Given this property, some of the very familiar statistics that we are used to, such as the mean (average) now becomes a relevant measure of centrality. If we are willing to assign interval data into categories (e.g., 1–5, 5–15, 16– and up) we can still use cross tabulation to check association between two variables. However, the new power inherent in the data allows a whole new series of analyses where we look at association between variables using some form of correlation (*see* PRODUCT MOMENT CORRELATION).

With interval data the difference between 22 and 27 is the same as the difference between 5 and 10. So, what do ratio data have that interval data do not? With ratio data, the value of 0 has a meaning in it that means nil, or nothing. Clearly, with the Fahrenheit scale a temperature of 0 does not mean 0 or nothing. The number of boxes of Cheerios consumed in a year would be a good example of a ratio scale. For all but the

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

most esoteric considerations such as a geometric or harmonic mean, ratio scales do not really provide much by way of practical value than interval scales.

Scales selected are important because they define the properties of the measurements that dictate the analyses that are possible subsequent to data collection. Where possible, marketers should try to collect interval data unless only ordinal data or nominal data fit the situation.

Much of marketing is all about association between variables, where we might then make causal inferences. Many techniques exist that relate mixes of nominal, ordinal, and interval scaled data. For example, if we have an intervally scaled dependent variable that we think is dependent on one or more nominally scaled independent variables, we might use MULTIPLE REGRESSION, a correlation technique, by converting nominally scaled data into dummy variables.

Bibliography

Churchill, G.A. and Iacobucci, D. (2005) *Marketing Research: Methodological Foundations*, Thompson South-Western, Mason.

Zikmund, W.G. and Babin, B.J. (2007) *Exploring Marketing Research*, Thompson South-Western, Mason.

probability sampling

Richard J. Fox

The purpose of sampling is to select for analysis a subset of the population to make inferences about the entire population (typically a large group). The population can be people, households, geographically defined areas, manufactured items in a warehouse awaiting shipping, and so forth. Measurements associated with items in the selected subset are used to develop estimates of descriptive statistics for the population such as the mean, mode, and standard deviation. Statistical assurance regarding the integrity of these estimates as well as the associated precision calculations is based on the assumption that probability sampling was used to select the sample. This is a crucial assumption, but it is often overlooked in practice.

Probability sampling refers to any method of selecting the subset of the population that comprises the sample in such a way that each element of the population has a positive probability of being included in the selected subset. Suppose the population contains N elements, and the sample contains $n < N$ elements. If the selection process is such that all subsets of size n are equally likely to be the sample selected, the process is called *simple random sampling*, a form of probability sampling. The probability of any element being selected in the sample is n/N. In most situations N is large, and even for relatively small values of n there are many possible subsets. Consequently, the number of possible estimates of any specific population characteristic of interest can also be very large.

Suppose an estimate of the average income of all N households in a specific geographic region is desired, based on a sample of size n. The sample average based on the n selected households in the sample is an estimate of the mean for the entire population. The value of this estimate varies from sample to sample. If simple random sampling is employed, then the average value of the individual sample means associated with the numerous possible samples is, in fact, equal to the population mean. In statistical language, the estimate (sample average) is said to be unbiased. Furthermore, if simple random sampling is used, one can readily calculate the 95% confidence interval (+/− error limits) from the sample data. There is only a 5% chance that the actual population mean is not contained in this 95% confidence interval, which provides insight into the precision of the estimate. These intervals can be calculated at alternative confidence levels (e.g., 90% or 99%). As the level of confidence increases, so does the width of the calculated interval. Also, as the sample size increases, the width of the confidence interval decreases for a fixed level of confidence (*see* STATISTICAL APPROACHES TO DETERMINING SAMPLE SIZES). Finally, the width of the confidence interval increases as the variance of the measurements associated with the sample elements increases.

There are other forms of probability sampling including *systematic sampling*, *cluster sampling*, and *stratified sampling*, (see Cochran (1977) for a thorough discussion of these sampling methods), which are more sophisticated than simple random sampling. The calculations of the estimators and the associated levels of precision vary with the type of probability sampling used, and can become quite complicated. These alternative forms of probability sampling can increase efficiency and cost effectiveness, and are sometimes preferable because of practical considerations (*see* SAMPLING TECHNIQUES).

In summary, it is important that probability sampling be employed when selecting a sample from which inferences regarding a population are to be made. Only then can statistical theory be used to guarantee integrity (unbiasedness) of the estimates and determine the precision of the associated estimates. Too often, samples are selected without using a probability sample, and statistical methods are still used to establish the precision of the subsequent estimates. Estimates derived from such samples can be useful, but traditional statistical precision calculations are not applicable.

Bibliography

Cochran, W.G. (1977) *Sampling Techniques*, 3rd edn, John Wiley & Sons, Inc. New York.

Thompson, S.K. (2002) *Sampling (Wiley Series in Probability and Statistics)*, John Wiley & Sons, Inc., New York.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

projective techniques

Linda L. Price

Projective techniques are characterized by substantial methodological heterogeneity that presents both research opportunities and challenges. Classically, projective techniques use stimuli to allow the research participant to articulate repressed or withheld feelings by projecting these onto another character or object. However, many researchers also group enabling techniques with projective techniques. Enabling techniques are devices that allow the research participant to find a means of expressing feelings, thoughts, and so on that they find hard to articulate. In marketing research, the distinction between projective and enabling techniques is often blurred and they are used similarly (*see* MOTIVATION RESEARCH). In this description they are combined.

Despite high variance, there are distinct similarities among projective techniques both in the nature of the research stimuli and the type of data elicited. Projective techniques typically rely on *indirect questioning*, and *ambiguous stimuli* that allow for nearly limitless variations in informant responses. For example, consumers might be asked to construct collages to represent their desires, describe dreams involving fast food, characterize a typical target shopper, or draw a "couch potato." Projective techniques can uncover *primary motives* as well as *latent motives* that consumers might otherwise be unwilling or unable to articulate. In addition, they tend to elicit data that compared with many other research approaches is more *sensory, emotional, symbolic, metaphorical*, *and imaginative*. Finally, projective techniques are relatively *free of social desirability bias* (*see* SOCIAL DESIRABILITY BIAS). The indirect questioning and ambiguous stimuli give informants the opportunity and permission to express their true and deeper level concerns.

Projective techniques have their historical roots in clinical psychology and have been the target of critical debate and scathing criticism as tools for patient diagnosis, treatment, and prescription because of justifiable concerns about their reliability and validity for this type of use. Partially in response to these kinds of concerns, projective techniques passed out of popular use in marketing and consumer research in the 1970s and 1980s. However, there has been a resurgence in attention and use encouraged by theoretical interests in consumer's emotions, desires, motivations, brand meanings, and relationships all of which are well suited to examination using projective techniques (*see* CONSUMER DESIRE; MOTIVATION AND GOALS). In market and consumer research, projective techniques are primarily used to discover new ideas, broaden customer understandings, and generate new market offerings, brand positions, or communications. They are more likely to be employed in theory building than testing and are rarely used exclusively, but most often in combination with other methods such as depth interviews or surveys (*see* FOCUS GROUPS AND DEPTH INTERVIEWS; SURVEY RESEARCH).

As with other primarily interpretive methods, the skill and experience of the researcher is paramount in constructing a trustworthy account. Several criteria can be employed to ensure the trustworthiness of the interpretation including adequacy, resonance, integration, innovation, and usefulness. To enhance reliability and validity, researchers may triangulate across different types of data or may elicit member checks from participants following researcher interpretations. Some argue that because projective techniques enable participants to express themselves in fuller, more subtle, and fairer ways (without the constraints of direct questioning) they may actually achieve greater validity than other methods (*see* RESEARCH RELIABILITY AND VALIDITY).

In deciding whether the use of projective techniques will add value, researchers should consider consumer memory, communication, and self-disclosure concerns (*see* CONSUMER MEMORY PROCESSES). For example, the use of story telling, dream exercises, autodriving with personal photos, and collage construction can help consumers recall and/or help them retrieve the emotionality or meaning of a consumer episode. Projective techniques that rely on nonverbal expressions such as picture drawing or collages can sometimes facilitate consumer articulation or encourage elaboration. And, of course, when consumers are reluctant to

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

disclose or inclined to lie, projective techniques may add significant value.

Projective techniques differ both in the quantity of data they elicit and in their response modality. With regard to quantity of data, techniques such as word association, sentence completion, or cartoon tests elicit relatively little raw data per informant while collage construction, dream exercises, autodriving, thematic stories, and psychodramas elicit a large amount of raw data per informant. Modality of response can vary from short written/verbal responses to lengthy verbal/written narratives, and from brief visual associations such as symbol matching to complex visual constructions such as collages and drawings. Responses may also take a physical/body language form such as when informants role-play in a psychodrama. By including projective techniques that vary in modality and quantity of response, researchers may be able to increase their interpretive opportunities.

Bibliography

Belk, R.W., Ger, G., and Askegaard, S. (2003) The fire of desire: a multisited inquiry into consumer passion. *Journal of Consumer Research*, **30** (3), 326–351.

Boddy, C. (2005) Projective techniques in market research: valueless subjectivity or insightful reality? *International Journal of Market Research*, **47** (3), 239–254.

Heisley, D.D. and Levy, S.J. (1991) Autodriving: a photoelicitation technique. *Journal of Consumer Research*, **18** (3), 257–272.

Levy, S.J. (1985) Dreams, fairy tales, animals and cars. *Psychology and Marketing*, **2** (2), 67–81.

Levy, S.J. (1996) Stalking the amphisbaena. *Journal of Consumer Research*, **23** (3), 163–176.

Mariampolski, H. (2005) *Ethnography for Marketers: A Guide to Consumer Immersion*, Sage Publications, Thousand Oaks.

Rook, D.W. (2006) Let's pretend: projective methods reconsidered, in *Handbook of Qualitative Research Methods in Marketing* (R.W. Belk), Edward Elgar Publishing, Inc., Northampton, pp. 143–155.

Spiggle, S. (1994) Analysis and interpretation of qualitative data in consumer research. *Journal of Consumer Research*, **21**, 491–503.

Zaltman, G. and Coulter, R. (1995) Seeing the voice of the customer: metaphor-based advertising research. *Journal of Advertising Research*, **35** (4), 35–51.

validity and reliability

Frederick Winter

When we take measurements we want them to mean something. Assume a person wants to get healthy, and therefore wants to lose weight. He or she would get pretty discouraged with his or her weight loss program if the bathroom scale is so unreliable that the weight fluctuates significantly from day to day – not because of gravitational pull (the person's true weight) but because the scale is erratic. We would call this the *issue of reliability*.

We quite often put consumers into experimental groups for analysis (*see* EXPERIMENTAL DESIGN) when some of this unreliability of measurement might seem tolerable. For example, one group of consumers eats vegetables and gets lots of exercise and the other group eats at McDonald's and watches TV all day. When we compare the two groups, even with an erratic scale, it is expected that the differences would tend to cancel out and we would still see the vegetable-exercise group make progress? Perhaps this may happen, but we must remember that significance tests generally look at the difference between two means divided by the variance, and, if the unreliable scale increases the variance, we are less likely to detect a significant difference when one might really exist. Marketers improve their reasoning ability from data when they use reliable measures.

Marketers are able to assess reliability in a number of ways. One method is the *test–retest reliability*. Measure something and then remeasure it again. When it is reasonable to assume nothing has changed, we should get close to the same measurement. The simplest and most common way is to calculate a simple correlation coefficient (range −1 to +1) between the two measurements (*see* PRODUCT MOMENT CORRELATION). High correlation suggests reliability.

Another method is *split half reliability* where half the measures that claim to measure the same attribute in one subsample are correlated with the other half of the measurements that measure the attribute in the other subsample. *Coefficient alpha* (Cronbach, 1951 – see the original reference for the more generalized formulation) is the correlation between the two halves. When alpha is high, the measures are reliable. In the split half case (two subsamples), alpha is calculated as

$$\alpha = 2\left[1 - \frac{(\sigma_1^2 + \sigma_2^2)}{\sigma_{1+2}^2}\right] \qquad (1)$$

where σ_x^2 is the variance of variance of subsample 1, 2 or the combined subsamples (1 + 2). Although the formula does not look intuitive, recall that the variance of a two-term sum is the variance of each component plus two times the covariance (correlation). If correlation between the subsample is 0, the last term will equal 1, and $\alpha = 0$. If correlation equals 1, the last term will equal $1/2$ and $\alpha = 1$.

What about validity? Marketing is unlike the physical sciences where there exist precise definitions of weight/mass, distance, and so on. We measure how much someone likes something, a construct, with operational definitions – measures that we take – perhaps on a 1–5 numerical scale. Using our first example, a health professional might say that weight relates to overall health, but it is not a perfect measure. This is the issue of validity. Is weight and weight loss a good indicator of health? Marketers generally consider face (or content) validity, criterion (predictive/concurrent) validity, and construct validity as ways in which to make sure that their measurements are valid.

Face validity or content validity is based on a subjective judgment as to whether it makes sense that what we are measuring indeed represents the underlying construct we are trying to measure. One rule of thumb is to observe how others have chosen to measure it and depart from this only if it seems to make great sense. Let us say we are trying to measure sales force effort – is it simply the number of calls? The number of calls really does not consider the territory size, distance that has to be traveled, difficulty of getting an appointment with the customer mix in the region, and so on. More than likely we might want to use a weighted combination of some of the above and that would likely result in more content validity.

Criterion validity is where the researcher has a theory with which he or she is pretty comfortable and which involves that which is being

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

measured. The researcher then makes sure there is correlation (concurrent validity if the measurements are expected to correspond at the same point in time and predictive validity if one measure is expected to relate to another at a later point in time). For example, if the measure of salesperson effort is known, the researcher would then correlate it with sales performance, since he or she is somewhat comfortable with the conclusion that more effort should result in greater performance. Of course, the problem is that other things might contribute to sales performance that the researcher would attempt to hold as constant – either statistically with some form of analysis of covariance or stepwise regression or perhaps by controlling with sampling. For example, experience could be held constant by simply testing whether the measure of sales effort predicted sales among salespersons in their fourth year of employment.

Construct validity is the most difficult to measure and less precise when one does. This is where the researcher makes sure that he or she is really measuring the construct that he or she reports to be measuring. Let us say part of our theory is that wealth affects purchasing behavior. If our measure of wealth has construct validity, then we might expect it to have criterion validity with some known other factors that we know wealth relates to (having an estate plan, using tax accounting services, value of real estate held, etc.) and it probably would be uncorrelated (discriminant validity) with some other measures that we know to be unrelated to wealth (it should be uncorrelated with happiness if it is well established that wealth does not relate to happiness). In short, does it fit as we would expect it to fit?

Reliability and validity interact to produce measurements in which we can have confidence – valid measures are reliable, but reliable measures are not necessarily valid (they might be reliably measuring another construct). However, if measures are not reliable, they cannot be valid.

Bibliography

Cronbach, L.J. (1951) Coefficient alpha and the internal structure of tests. *Psychometrika*, **16** (3), 297–334.

Churchill, G.A. and Iacobucci, D. (2005) *Marketing Research: Methodological Foundations*, Thompson South-Western, Mason.

Zikmund, W.G. and Babin, B.J. (2007) *Exploring Marketing Research*, Thompson South-Western, Mason.

social desirability bias

Pamela Grimm

Social desirability bias refers to the tendency of research subjects to choose responses they believe are more socially desirable or acceptable rather than choosing responses that are reflective of their true thoughts or feelings. This tendency results in over-reporting of responses that are socially desirable and under-reporting of those responses that are deemed to be socially undesirable/less desirable. This tendency is believed to be a personality trait based on the subject's need for approval. Social desirability bias may become an issue when research involves collecting data on personal or socially sensitive issues. For example, respondents may believe that society expects them to be responsible for the environment, even though the respondent may not care much about the environment. In such a situation, the respondent may simply indicate an environmentally conscious attitude rather than choose an option that represents his or her true feelings. The idea of "political correctness" is based on social desirability bias.

The problem of social desirability bias is most likely to occur in those situations in which questions relate to what are widely accepted attitudes, or behavioral or social norms (e.g., smoking, drug use, lying, cheating, and protection of environment). This bias is reported to be most apparent when data are collected through a survey method where the respondent can be easily identified. For example, social desirability bias may occur when collecting data through personal interviews, specifically because of the presence of another individual. Collecting data through a mechanism that allows for privacy may decrease social desirability bias. For example, a mail survey may be least susceptible to such bias because of the impersonal nature of responses and the ability of the respondent to submit the survey anonymously. Reassuring subjects that their responses will be kept confidential or anonymous should also help minimize social desirability bias.

The concept of social desirability bias is also related to a concept termed *demand effects*. The idea behind demand effects is that the subject responds to a question in the way he or she believes the researcher desires. Demand effects are specific to the researcher and may or may not reference the subject's beliefs about what society as a whole believes about a specific topic. Careful design of a survey or focus group outline, which avoids leading questions, should be able to minimize or reduce demand effects.

Recent research on social desirability bias indicates that this bias may not be as prevalent as previously believed. Part of "intentional misrepresentation" may actually be attributable to accidental mistakes in recall (Krosnick, 1999, p. 546), rather than intentional lying.

Researchers have suggested a number of actions to address this issue (Nederhof, 1985). If the length of the survey permits, it is advisable to measure the social desirability bias by administering the social desirability scale. This allows the researcher to assess the extent of the problem of social desirability influence in survey responses. A properly designed survey could also help mitigate this problem. For example, in the case of questions that are prone to generate "socially desirable" responses, it is important to provide response options to such questions where the respondents have the opportunity to tell the truth. One of the sources of social desirability influence is the interviewer himself/herself, whereby the respondents try to "please" the interviewer by guessing what are "acceptable answers." When possible, collecting data through mail surveys rather than telephone or face-to-face surveys tends to reduce the social desirability bias. If it is necessary to use other modes of data collection using an interviewer, then it is advisable to use professional interviewers.

Bibliography

Krosnick, J.A. (1999) Survey methodology. *Annual Review of Psychology*, **50**, 537–567.

Nederhof, A.J. (1985) Methods of coping with social desirability bias: a review. *European Journal of Social Psychology*, **15**, 263–280.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

concept of causality and conditions for causality

Harmen Oppewal

Causality concerns relationships between variables where a change in one variable necessarily results in a change in another variable. Philosophers have debated for centuries whether there is such a thing as causality. Can we ever know the true causes of anything? Empiricists like the eighteenth-century Scottish philosopher David Hume argued that all we can see are the observable associations between events (Dunbar, 1995). Causes are only in our minds and we can never be sure they really exist. Nevertheless, causal knowledge is one of our most powerful resources because it tells what can or should be done to obtain a desired consequence or to avoid an undesirable outcome (Shadish, Cook, and Campbell, 2002). Causal knowledge is especially useful if it includes knowledge about the effects of variables that we have active control over – in marketing these are the typical marketing mix variables such as price, promotion, place (location and distribution), and, of course, the actual product.

Three Requirements

There are *three requirements* for a causal relationship, according to John Stuart Mill, a nineteenth-century English philosopher. In general, as shown in Figure 1, the question is when is X a cause for Y? Firstly, there has to be covariation: if X varies then Y should vary too. If there are clouds (X), then it may start raining, but if there are no clouds, it should not rain – or we would be very surprised, as it violates a basic piece of causal knowledge that is well ingrained in our shared understanding of how things work in our world: clouds are a prerequisite for rain. Similarly, in marketing we generally expect promotions to have at least some effect on sales and would be surprised (if not disappointed) if variations in promotional activities did not correspond with at least some variation in sales.

Secondly, there should be a sequence in time: First X changes, and then Y changes. We would be very suspicious if it first rained and clouds would appear only later. For promotions and sales levels, this relationship is more complicated as many different factors drive sales. Our challenge is to disentangle the variation due to other factors from the variation that we expect to occur due to promotion; once we achieve this, we expect variation due to promotional activity only to occur after the campaign has started. In contrast, if Y changes without X changing or before X has changed, we must conclude there is no causal relationship.

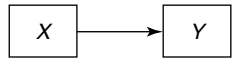


Figure 1 A true causal relationship between X and Y.

The third requirement is the most difficult to determine: No other, or "third variable," should covary with X and Y. A "*third variable*" is a variable, possibly yet unidentified or unobserved, that might be the true cause for changes observed in two other variables of interest. If there were a variable Z that varies with X and Y, Z instead of X could be the true cause for the variation observed for Y. This relationship is shown in Figure 2. In the rain example, this would mean that, whereas we may have recorded heavy clouds overnight, we cannot be sure the increased water levels in a stream are due to rain or are the result of melting snow upstream. Similarly, there may be other factors driving increases in sales at the same time as we introduced our campaign. Although the sales increase coincides with the introduction of the campaign, it is well possible that the campaign has no real effect because the increase is caused by other factors such as a change in competitor activity or a seasonal effect like the start of the holidays. In this case, if the campaign content is linked to a season like Christmas or back to school, there is a double causal relation between the season and the change in the studied "cause" and the observed effect. In this case, the relationship between having a promotional campaign (X) and observed sales (Y) is completely *spurious* because both variables directly depend on the start of the holiday season (Z) and a conclusion that the campaign causes sales to increase is false.

Alternatively, the true cause may indeed be the promotional campaign, but this leaves the question unanswered as to what was the "active

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

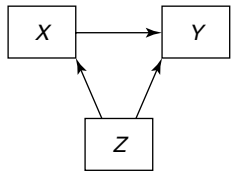


Figure 2 A spurious relationship between X and Y.

ingredient" that made the promotion a success. Was it the media coverage, or the type of promotion, or the role of imagery or type of appeal? Or were these "ingredients" all strictly required to make the promotion achieve the observed sales effect?

Conclusion

In sum, causal knowledge is valuable but difficult to attain. Not only are there alternative explanations, often we are not sure what exactly the "active ingredient" is in our observed causes. This is a crucial but often underestimated issue in applied research. Indeed, too often researchers spend a lot of time developing, implementing, and analyzing a design without careful consideration of the essential question: what do we think we are studying and does it match what we are actually studying? This is the question of construct validity (*see* VALIDITY IN EXPERIMENTATION; RESEARCH RELIABILITY AND VALIDITY).

Bibliography

Dunbar, R. (1995) *The Trouble with Science*, Faber and Faber, London.

Shadish, W.R., Cook, Th.D., and Campbell, D.T. (2002) *Experimental and Quasi Experimental Designs for Generalized Causal Inference*, Houghton Mifflin, Boston.

confirmatory factor analysis

Richard J. Fox

Introduction

The basic factor analysis model is a representation of relationships between a set of observed variables, all of which are manifestations of a smaller set of unobserved latent variables or factors. Each observed variable is a linear combination of the unobserved latent factors, which may or may not be correlated, plus a random error term, also unobserved. The random unobserved error terms can also be correlated, but are independent of the unobserved latent factors. The coefficients of the latent factors are called *factor loadings* and they reflect the degree of association between the corresponding latent and observed variables. The matrix equation displaying the covariance matrix of the observed variables (left-hand side of the equation) as a function of three matrices, the matrix of factor loadings, and the covariance matrices of the latent variables and the random error terms, is called the *covariance equation*.

The model has a large number of unknown parameters, all elements of the three matrices appearing in the right-hand side of the covariance equation, and constraints must be imposed to reduce the number of parameters to be estimated. Selected pairs of latent variables or random error terms can be assumed to be independent (zero covariance). Each latent variable may be assumed to manifest itself in only a specified subset of the observed variables (i.e., selected factor loadings are set equal to zero). Also, equality of variance constraints may be imposed among the errors or latent factors.

Identification

Constraints are imposed for the purpose of model identification (see Long's *Confirmatory Factor Analysis* for a discussion of model identification), which means that a unique set of parameter estimates will be obtained for the constrained model. If a model is unidentified, an infinite set of possible model parameter estimates produce the same covariance matrix for the observed variables. If the model is identified, estimation can proceed. A random sample of the observable variables is used to develop a sample covariance matrix, and maximum likelihood methods are typically used to estimate the model matrices (the three matrices described above) under the imposed constraints. The estimates of the model matrices are chosen so that the matrix formed by applying the covariance equation, using the estimated matrices, produces a covariance matrix that is as "close" as possible to the sample covariance matrix. Finally, the estimated covariance matrix corresponding to the constrained model is compared to the sample covariance matrix to determine whether the structure imposed by the model assumptions and constraints is consistent with the data. The chi-square goodness-of-fit test is used to compare these two matrices. The larger the difference, the larger the chi-square statistic. A relatively "small" chi-square "confirms" the structure resulting from the model constraints; a relatively "large" chi-square value leads to rejection of the assumed model. This is the essence of confirmatory factor analysis.

Example

Consider as an example, the SERVQUAL Scale, developed by Zeithaml, Parasuraman, and Berry (1990) designed to measure the gap between customer quality expectations and perceived delivery of key service dimensions. The RATER version of SERVQUAL focuses on five dimensions of service quality: reliability, assurance, tangibles, empathy, and responsiveness. To compute a SERVQUAL score, survey participants (customers) first rate their levels of agreement for a list of over 20 statements reflecting expectations related to the five core service dimensions. (The rating exercise is repeated in the context of service provider performance delivery, and the gap is measured by comparing expectations to delivery). The set of over 20 item ratings constitutes the set of observed variables, and the latent factors are the five service dimensions. Early development work associated with SERVQUAL was designed to investigate the relationships between the list of statements and core service dimensions. For example, can the list of statements be reasonably

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

partitioned into five mutually exclusive and exhaustive sets? If each statement in the list can be associated with one, and only one, of the core dimensions, then each dimension can be viewed as an aggregate of a subset of list items. This question can be addressed using confirmatory factor analysis.

In the SERVQUAL example, the use of five mutually exclusive and exhaustive sets of statements reflecting five core service dimensions implies that many of the factor loadings are zero. In addition, the five latent service factors can be assumed to be independent, and so on. A sample covariance matrix is calculated from a random sample of respondents' ratings of the list of over 20 items. The three matrices of the covariance equation, suggested by the assumed data structure, are estimated using maximum likelihood applied to the same random sample of respondents ratings. The estimated covariance matrix obtained by applying the covariance equation using these three estimated matrices is then compared to the sample covariance matrix. If the two covariance matrices are "close," the hypothesized data structure for the observed variables (ratings) is consistent with the data.

Structural Equation Models

Confirmatory factor analysis is also often a critical component of what is referred to as *structural equations modeling*. In this type of model building, a causal relationship is presumed to exist among a set of latent factors. These latent factors manifest themselves through a set of variables that are possibly measured with error. Confirmatory factor analysis is used to develop a measurement model for the unobserved latent variables, which are the basic variables in the causal relationship being explored.

Exploratory Factor Analysis

Confirmatory factor analysis might also be used in conjunction with exploratory factor analysis (see Applied Multivariate Research Design and Interpretation by Meyers, Gamst, and Guarino (2006) for a discussion of exploratory versus confirmatory factor analysis). For example, exploratory factor analysis might be applied to half of the sample data to develop a model relating latent factors to observed variables. The exploratory analysis would determine how many underlying latent factors there are, as well as which factors influence which observed variables. Confirmatory factor analysis can then be used on the other half of the sample to determine how well the model fits this independent data set. The observed covariance matrix for this "hold-out" sample would be used to assess how well the structure imposed by the model derived by exploratory factor analysis fits an independent data set.

Bibliography

Long, J.S. (1983) *Confirmatory Factor Analysis*, Sage Publications, Inc., Thousand Oaks.

Mayers, L.S., Gamst, G., and Guarino, A.J. (2006) *Applied Multivariate Research Design and Interpretation*, Sage Publication, Inc., Thousand Oaks.

Parasuraman, A., Zeithaml, V.A., and Berry, L.L. (1988) SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, **64** (1) 12–40.

Savalei, V. and Bentler, P.M. (2006) Structural equations modeling, in *The Handbook of Marketing Research* (R. Grover and M. Veiens), Sage Publications, Inc., Thousand Oaks.

Zeithaml, V., Parasuraman, A., and Berry, L.L. (1990) *Delivering Quality Service; Balancing Customer Perceptions and Expectations*, Free Press, New York.

content analysis

Linda L. Price

Content analysis is a general term for techniques that *systematically* identify and evaluate the symbolic content of communications (images, words, etc.). It can be effectively employed to examine virtually any type of communication and may focus on either quantitative or qualitative aspects of messages. While some interpretive researchers dismiss content analysis from discussion of qualitative methods emphasizing it is quantitative, others argue for expanding the term and its application to include accounts of communications not in the form of numbers and not restricted to frequency counts. Quantitatively, researchers can create tally sheets to determine frequencies of relevant categories in the way that content analysis is typically employed. Alternatively, researchers can qualitatively examine themes, topics, symbols, and other phenomena while systematically grounding such examinations in the data.

Content analysis of text is distinct from a hermeneutical approach to analysis in that data are analyzed by means of explicit rules called *criteria of selection* and these are formally established *before* the actual analysis of the data (*see* ETHNOGRAPIC RESEARCH). The researcher seeks to establish criteria for selection that are sufficiently exhaustive that they can account for each variation in communication content and then apply these criteria such that others looking at the same messages would obtain comparable results. Although content analysis is often viewed as objective because of an emphasis on agreement between coders (often naïve to research aims or hypotheses), the researcher still makes many interpretive judgments. This is especially true if codes move beyond the manifest content (the actual words or images that can be counted) to more latent or symbolic meanings conveyed by a message. Arguably, most categories of interest to market and consumer researchers are open to at least some interpretive ambiguity. For example, if a researcher wants to count instances of a theme, she must first establish what strings of words or images are constitutive of that theme. Additionally, even with simple word or character counts researchers must be careful about what inferences they draw from descriptive statistics such as proportions and frequency distributions. For example, just because a buyer uses twice as many positive words in a product review as someone else, does not mean he is twice as happy with his purchase.

Content analysis is useful for researchers attempting to identify patterns, frequencies, or potential categories with respect to marketing and consumer phenomena. As used in marketing and consumer research, it is primarily a descriptive technique (*see* DESCRIPTIVE RESEARCH). Content analysis offers a number of benefits. It allows for an unobtrusive appraisal of communications; it can provide an empirical starting point for building theory about the nature and effect of specific communications; it can link message content to environmental variables, source characteristics and so on, and it can be effectively integrated with other types of qualitative and quantitative analyses.

The application of content analysis to marketing problems is likely to increase in the next decade. With an explosion of data in digital form, computer-assisted content analyses can now provide fine-grained semantic analysis of thousands of communications daily. One area that has received a lot of attention is *automatic sentiment analysis* which classifies documents, or chunks of text, into emotive categories such as positive or negative. There are numerous problems in marketing and consumer research that might be addressed by exploring the emotional content of large collections of open domain documents. For example, a company might use sentiment analysis to understand customer satisfaction by examining the range of affect in product reviews, what products are viewed most positively or negatively, how affect is linked to different features of a product and so on (*see* CUSTOMER-SATISFACTION RESEARCH). There are currently two basic approaches for identifying sentiment in text. Text classification models first hand label a set of documents for affect and then a system is trained on the vector features associated with the vector set. In the lexical approach, documents, phrases, or

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

sentences are categorized as positive or negative on the basis of their match to a lexicon of sentiment-bearing terms. There are many challenges in doing this type of research but the burgeoning applications of these analytical tools suggests that this will be a high-growth application of content analysis in the future.

Bibliography

Ahuvia, A. (2001) Traditional, interpretive and reception based content analyses: improving the ability of content analysis to address issues of pragmatic and theoretical concern. *Social Indicators Research*, **54**, 139–172.

Berg, B.L. (1998) *Qualitative Research Methods in the Social Sciences*, 3rd edn, Allyn & Bacon, Needham Heights.

Carlson, L. (2008) Use, misuse, and abuse of content analysis for research on the consumer interest. *Journal of Consumer Affairs*, **42** (1), 100–105.

Gregory, M.L., Chinchor, N., Whitney, P. *et al*. (2006) *User-directed Sentiment Analysis: Visualizing the Affective Content of Documents*. Proceedings of the Workshop on Sentiment and Subjectivity in Text, July, 23–30, Sydney.

Kolbe, R.H. and Burnett, M. (1991) Content-analysis research: an examination of applications with directives for improving research reliability and objectivity. *Journal of Consumer Research*, **18**, 243–250.

Krippendorff, K. (2003) *Content Analysis: An Introduction to its Methodology*, 2nd edn, Sage Publications, Inc, Thousand Oaks.

West, M.D. (2001) *Theory, Method, and Practice in Computer Content Analysis*, Ablex Publishing, Westport.

cross-tabulation

Mayukh Dass

When we are interested to determine how two or more categorical variables are linked in a study, cross-tabulation is one of the initial approaches to consider. In a cross-tabulation, we create a table similar to that of a FREQUENCY DISTRIBUTION to learn more about frequency distribution, but here, we merge the counts of different values of two or more variables. In other words, a cross-tabulation table is a way to present frequency distribution of two or more variables concurrently. The different values of the variables are considered categories, and they are cross-classified among each other (Malhotra, 2007). The cross-tabulation table is also known as a *contingency table*.

Although cross-tabulation tables can be generated for any number of variables, most useful tables for managerial purposes are formed with two or three variables. Table 1 illustrates an example of a two-variable cross-tabulation, which is also known as *bivariate cross-tabulation*. In this example, students from an undergraduate market research class were asked about their liking toward pink lemonade. Their responses along with their gender identification are shown in Table 1. Since the two variables "likes pink lemonade" and "gender" are cross-classified, we can also compute the row and column percentages for each of the values. From the manager's perspective, such a bivariate cross-tabulation provides information on (i) how the values of the two variables are related, (ii) which cross-classification is most selected by the respondents, and (iii) how these cross-classifications are different from each other. From the example shown in Table 1, we find that female respondents liked pink lemonade more than male respondents. More precisely, 85.71% (= 30/35) of the female respondents like a pink lemonade as compared to 68.57% of the male respondents.

When a bivariate cross-tabulation needs further clarification, a third variable, which is typically a categorization of one of the two variables, is considered. Continuing with the above example, let us include another variable, "gender age group," into the cross-tabulation to further explain the relationship between male and female respondents. The resulting three-variable cross-tabulation table (Table 2) provides greater

Table 1 Cross-tabulation of liking pink lemonade by gender.

Question: Do you like pink lemonade?		*Gender*		
		Male	*Female*	*Total*
Likes pink lemonade	Yes	24	30	54
	No	11	5	16
	Total	35	35	70

Table 2 Cross-tabulation of liking pink lemonade by gender by age group.

Question: Do you like pink lemonade?		*Gender*				
		Male Age Group		*Female Age Group*		*Total*
		Less than 21 years	*More than 21 years*	*Less than 21 years*	*More than 21 years*	
Likes pink lemonade	Yes	9	15	24	6	54
	No	4	7	3	2	16
	Total	19	16	27	8	70

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

details about the responses from two age-related subgroups of the variable "gender," that is, "less than 21 yrs old" and "more than 21 yrs old." Here, we find that pink lemonade is more popular among the female respondents who are less than 21 years than any other gender age groups. For further analysis on the relationship between the variables, NONPARAMETRIC TEST to learn more about nonparametric tests such as a two-sample chi-square test are performed on the cross-tabulation data.

Bibliography

Malhotra, N.K. (2007) *Marketing Research: An Applied Orientation*, 5th edn, Pearson Education, Inc., Upper Saddle River.

endogeneity

Neeraj Arora

Consider the regression model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$ (*see* MULTIPLE REGRESSION). To consistently estimate model parameter β_1 it is required that $E(\varepsilon) = 0$ and Cov $(x_1, \varepsilon) = 0$. That is, the expected value of the error distribution is zero and it is uncorrelated with variable x_1 included in the model. However, if Cov $(x_1, \varepsilon) \neq 0$, then x_1 is considered endogenous. Failure to recognize endogeneity in a model leads to a biased estimate for β and therefore, requires careful attention. Endogeneity could be caused by three factors (Wooldridge, 2002, p. 51):

- *Omitted variables*: Factors unobserved by the researcher, often referred to as *demand shocks*, can affect both the demand and the marketing mix variables set by a manager. For example (Villas–Boas and Winer 1999) observe that it is plausible that an unobserved variable such as a manufacturer coupon drop could affect a retailer's price that is observed and included in the model. This induces a correlation between the error term and the included price variable thus causing an endogeneity bias.
- *Measurement error*: Measures in a survey have an associated measurement error. If this error is correlated in a systematic manner with the model error term we have an endogeneity problem (Gilbride, Yang, and Allenby, 2004).
- *Simultaneity*: The selected level of an independent variable is sometimes a function of the dependent variable. For example, Manchanda, Rossi, and Chintagunta (2004) note that promotion decisions (x) are often made simultaneously with y – high potential markets often get greater resources. The variable x therefore is not exogenously determined.

To address endogeneity, one possible solution is the use of *instrument variable*s (IVs). This is accomplished by selecting a variable z_1 such that Cov$(z_1,\varepsilon) = 0$. That is, z_1 is exogenous to the system. Further z_1 and x_1 should be partially correlated. For example, it is not uncommon to use lagged price terms or prices from "similar" geographic locations as possible IVs for price. Marketing applications of this approach include those by Hui (2004) and Nair, and by Chintagunta, and Dubé (2004). The basic idea is to use a two-stage approach by letting the IV z_1 predict the independent variable x_1 and using this predicted x_1 as a regressor in the demand equation. A problem here is that selecting an IV is more an art than a science. It requires instruments to be independent of an unobserved quantity (ε) a priori (Chandukala *et al.*, 2008). For most marketing problems, this does not appear to be an attractive solution. Another approach is to recognize heuristics used by managers in setting marketing mix decisions. That is, the demand and marketing variables could be modeled jointly by recognizing that the x decisions may be explicitly related to the expected demand or model parameters. A third solution is to assume an objective function of the decision maker (Yang, Chen, and Allenby, 2003) that helps a manager set the marketing mix variables. The objective function includes the model parameters that are a part of the demand equation thus leading to a system of equations that need to be jointly modeled.

In marketing, there is an increased recognition of the limitations of a reduced form model that ignores the underlying data-generating mechanism. In fact, an overwhelming majority of academic and practitioner marketing research suffers from the reduced form criticism. This has led to greater efforts to use structural models (Franses, 2005; Chintagunta *et al.*, 2007) that carefully account for endogeneity problems, among others. Shugan (2004) argues that two more considerations with regard to endogeneity deserve a special mention. First, there is no substitute for better data. When possible, most relevant independent variables should be included. Data fusion from disparate sources such as print, TV, and radio should be collectively used to explain demand. This is an attempt to hit the omitted variable problem head-on. Second, we as marketers should conduct more experiments (*see* CONCEPT OF CAUSALITY AND CONDITIONS FOR CAUSALITY). Such

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

experiments should try to assess response sensitivities in an attempt to address simultaneity concerns.

Bibliography

Chandukala, S.R., Kim, J., Otter, T., Rossi, P.E. and Allenby, G.M. (2007) Choice models in marketing: economic assumptions, challenges, and trends. *Foundations and Trends in Marketing*, **2** (2), 97–184.

Chintagunta, P., Erdem, T., Rossi, P.E., and Wedel, M. (2007) Structural modeling in marketing: review and assessment. *Marketing Science*, **25** (6), 604–616.

Franses, P.H. (2005) On the use of econometric models for policy simulation in marketing. *Journal of Marketing Research*, **42** (1), 4–13.

Gilbride, T., Yang, S., and Allenby, G.M. (2004) Modeling simultaneity in survey data. *Quantitative Marketing and Economics*, **3**, 311–335.

Hui, K. (2004) Product variety under brand influence: an empirical investigation of personal computer demand. *Management Science*, **50** (5), 686–700.

Manchanda, P., Rossi, P., and Chintagunta, P. (2004) Response modeling with non-random marketing mix variable. *Journal of Marketing Research*, **41**, 467–478.

Nair, H., Chintagunta, P., and Dubé, J.P. (2004) Empirical analysis of indirect network effects in the market for personal digital assistants. *Quantitative Marketing and Economics*, **2** (1), 23–58.

Shugan, S. (2004) Editorial: endogeneity in marketing decision models. *Marketing Science*, **23** (10), 1–3.

Villas-Boas, J.M. and Winer, R. (1999) Endogeneity in brand choice models. *Management Science*, **45** (10), 1324–1338.

Wooldridge, J.M. (2002) *Econometric Analysis of Cross Section and Panel Data*, The MIT Press.

Yang, S., Chen, Y., and Allenby, G.M. (2003) Bayesian analysis of simultaneous demand and supply. *Quantitative Marketing and Economics*, **1**, 251–275.

frequency distribution

Mayukh Dass

In a frequency distribution, each variable in the study is considered separately and a response count along with their percentages and cumulative percentages are computed for all different values of the categorical variable. This information is presented in a table, termed as *frequency table*, as shown in Table 1. A frequency table is not only effective to determine how the respondents have answered the questions but it is also useful in identifying cases of nonresponses or illegitimate responses in the dataset.

In order to illustrate the effectiveness of this analysis, let us consider a frequency table (Table 1) taken from a study performed on art collectors and dealers who frequently participate in online auctions of contemporary Indian art. In this study, the respondents were asked questions related to their browsing and bidding activities on a 7-point scale, with 1 = strongly disagree and 7 = strongly agree. Here, we consider one of those questions that ask the respondents about their browsing behavior during the auction. The frequency table suggests that 80.5% of the respondents, or 33 out of total 41, agreed (*somewhat agree*, *agree*, and *strongly agree*) that they browse the auction web pages and see who is bidding on what items while participating in an online auction. As you can see, the frequency table is very informative for a quick assertion on the responses in a study. Apart from analyzing the responses, the frequency table can be further used to perform NONPARAMETRIC TEST to learn more about nonparametric tests such as chi-square tests to determine the distribution of the variable reported.

As mentioned earlier, frequency distribution is also useful in detecting data errors and for finding ways to recode the data. For example, let us consider another frequency distribution taken from the same study performed on art collectors and dealers of contemporary Indian art. Here, we asked the respondents about the number of paintings they purchase each year. The frequency table as shown in Table 2 suggests three interesting characteristics about the dataset. First, it suggests that most people purchase at least 1 painting (mode = 1 with 18 respondents) each year. Second, it shows that there is an error in the dataset as the table shows one response with −4 paintings, which is not possible in the given context. When such errors are detected by the frequency distribution table, one should cross-check the responses with the actual study instrument to determine the origin of the error and recode the record accordingly. In this dataset, we found that it was a coding error and the response was 4 paintings and not −4. And third, the frequency table shows that the distribution becomes thin or the number of responses become small after 3 paintings. This indicates that we can group responses with

Table 1 Frequency distribution of online bidder behavior.

Question: I Typically browse through the Web pages and see who is bidding on what items	*Frequency*	*Percentage*	*Cumulative Percentage*
Strongly disagree	0	0	0
Disagree	1	2.4	2.4
Somewhat disagree	3	7.3	9.8
Neither agree nor disagree	4	9.8	19.5
Somewhat agree	8	19.5	39.0
Agree	15	36.6	75.6
Strongly agree	10	24.4	100.0
N	41		

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 2 Frequency distribution of purchase behavior.

Question: How many paintings do you purchase each year?	*Frequency*	*Percentage*	*Cumulative Percentage*
−4	1	2.4	2.4
1	18	43.9	46.3
2	8	19.5	65.8
3	7	17.1	82.9
4	2	4.9	87.8
5	3	7.3	95.3
6	1	2.4	97.6
9	1	2.4	100.0
N	41		

Table 3 Frequency distribution of purchase behavior after recoding data.

Question: How many paintings do you purchase each year?	*Frequency*	*Percentage*	*Cumulative Percentage*
1	18	43.9	43.9
2	8	19.5	63.4
3	7	17.1	80.5
4 or more	8	19.5	100.0
N	41		

purchase of 4 paintings or more and recode the data accordingly to convert the continuous variable as a categorical variable and make it a more useful bidder behavioral factor in the study. A frequency distribution table after recoding the data is shown in Table 3.

Bibliography

Malhotra, N.K. (2009) *Marketing Research: An Applied Orientation*, 6th edn, Pearson Education, Inc., Upper Saddle River, NJ.

multicollinearity

George R. Franke

Predictor variables in MULTIPLE REGRESSION that are highly correlated provide little independent explanatory ability. A scatterplot of one such variable versus another will show points falling close to a straight line, hence the term *multicollinearity* or simply *collinearity*. Because collinearity can have serious consequences for regression analyses, an extensive literature has developed on ways to detect and treat it in empirical research (Belsley, Kuh, and Welsch, 1980).

Consequences and Indicators of Harmful Collinearity

Perfect collinearity results when two or more predictors are completely redundant. In such cases, the model cannot be analyzed as specified. Much more common is near collinearity, where the shared variance between predictors increases the variance of the estimated regression coefficients. Potential consequences of this increased variance include estimated coefficients that (i) have theoretically implausible magnitudes or signs; (ii) vary substantially with small changes in the sample of observations or the set of predictors; and (iii) are individually nonsignificant even though they explain significant amounts of variance as a group. These results may not reduce the model's ability to predict values of the criterion variable. They are much more serious if the goal is to understand the relative importance of the predictor variables (*see* STEPWISE REGRESSION).

When regression outcomes are not consistent with expectations, researchers may (rightly or wrongly) interpret the results as consequences of collinearity. Several more direct tests are available:

1. Correlations of above 0.8 or 0.9 between predictor variables are often interpreted as excessively collinear. However, if one of two highly correlated predictors has only a small correlation with the criterion variable, it can "suppress" irrelevant variance in the other predictor. This pattern boosts the variance explained by the model and reduces the standard errors of the regression coefficients. Similar results may occur if two predictors are positively correlated with the criterion but negatively correlated with each other (Mela and Kopalle, 2002). Therefore, researchers must examine the pattern and signs of the correlations and not just their magnitudes.
2. Sets of variables may be strongly related even though no two have a high pairwise correlation. Regressing each predictor x_i on all the other predictors gives the variance explained R_i^2. The inverse of the unexplained variance for x_i, $(1 - R_i^2)^{-1}$, is known as the variance inflation factor (VIF). The VIF indicates the increase in the variance of the regression coefficient for x_i that is due to collinearity. A VIF $\geq$10 is often thought to indicate harmful collinearity, though problems are possible with lower VIF values.
3. Condition indices (CI) are functions of eigenvalues of the scaled predictor variables. There are as many CIs as predictors, including the intercept, and each accounts for some proportion of the variance of each regression coefficient. Two or more predictors (including the intercept) that have large variance proportions, roughly 0.5 or greater, and that are associated with the same high condition index, larger than 20 or 30, may cause harmful collinearity.

Remedies for Harmful Collinearity

Given an existing data set, researchers may combine or omit predictors, or constrain coefficients in theoretically justifiable ways (e.g., to be equal). The benefits and disadvantages of these approaches depend on how well they match reality. Ridge regression provides biased (conservative) estimates of the regression coefficients, in exchange for lower standard errors. Therefore, the analysis potentially offers more precise estimates that are less affected by small changes in the data. STRUCTURAL EQUATION MODELING may deal with collinearity by examining causal paths through a sequence of correlated variables, rather than testing each one controlling for the others, or by treating highly

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

correlated variables as indicators of a common underlying construct.

Collinearity creates fewer problems when sample sizes are large, predictors have high variances relative to that of the criterion variable, and the model explains substantial variance in the criterion. Therefore, in many cases, appropriate model development and data collection procedures will be the best remedy for collinearity problems.

Bibliography

Belsley, D.A., Kuh, E., and Welsch, R.E. (1980, reprinted 2004) *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, John Wiley & Sons, Inc., Hoboken.

Mela, C.F. and Kopalle, P.K. (2002) The impact of collinearity on regression analysis: the asymmetric effect of negative and positive correlations. *Applied Economics*, **34**, 227–277.

nonparametric test

Mayukh Dass

In marketing research, when a study contains variables that are measured in interval or ratio scales (*see* ITEMIZED RATING SCALES (LIKERT, SEMANTIC DIFFERENTIAL, AND STAPEL)), their means, variances, and standard deviations are meaningful and can easily be computed. For example, the mean of respondents' propensity to purchase a brand measured on a seven-point scale, with 1 = highly unlikely, 7 = highly likely, will provide insights to the sample's desire to purchase the product. Here, the data can be assumed to be drawn from a particular probability distribution and can be put through various parametric tests (*see* HYPOTHESIS TESTING RELATED TO DIFFERENCES – PARAMETRIC TESTS). There are also other instances where the scales used in a study are either nominal or ordinal in nature, and the measures have no clear numeric interpretation. For example, in case of gender classification of customers, "male" or "female" (nominal scale) are just used as a labeling scheme. Similarly, when customers provide a preference ranking of different brands of breakfast cereal, their responses (ordinal scale) include an order in which the brands are preferred, but not the difference in preference between them. In such situations, it is inappropriate to compute means, variances, and standard deviations of the variables as they do not provide any meaningful insights. Further, the assumption that the data are drawn from a particular distribution is not valid. In such cases, a different set of analyses called *nonparametric analyses* can only be used to analyze the data. The tests used for hypotheses testing are called *nonparametric tests*.

Nonparametric tests are performed when the variables measured are either in nominal or in ordinal scale. We can also use these tests on variables measured in interval or ratio scale if we cannot make any assumptions about the population probability distribution (Conover, 1999). Like parametric tests, nonparametric tests are categorized into one- and two-sample tests. Two-sample tests are further classified into independent and paired-samples tests. While dealing with only one sample, we can perform five nonparametric tests: the Chi-square test, the Kolmogorov–Smirnov (K–S) test, the runs test, binomial test, and the sign test. The first four tests are different types of goodness-of-fit tests where the *Chi-square test* investigates whether the data came from a particular distribution, the *K–S test* compares the cumulative distribution functions for a variable with a specified distribution, the *runs test* examines the degree of randomness for dichotomous variables, and the *binomial test* investigates whether the observed responses for each category of a dichotomous variable follows a specified binomial distribution. Finally, the *sign test* is used to determine whether there are significant differences in preference between the two items of a dichotomous variable (Anderson, Sweeney and Williams, 2008).

In studies where two independent samples of nonmetric data are considered, four nonparametric tests are used: the *Chi-Square test*, the *Mann–Whitney U test*, the *median*, and *the K–S* test. In the given context, the *chi-square test* and the *K–S test* determine whether the two samples have the same probability distribution, the *Mann–Whitney U test* compares the difference in the location of two populations on the basis of the observations from the two samples and the *median test* examines whether the two populations from where the two samples are drawn have the same median. For paired samples, we can also perform four nonparametric tests including *sign test*, *Wilcoxon signed-rank test*, *McNemar test*, and *Chi-square test*. The *sign test* determines whether the paired differences of the paired sample is above or below a particular value, and the *Wilcoxon signed-rank test*, the *McNemar test*, and the *Chi-square test* examine the differences between the paired observations.

Bibliography

Anderson, D.R., Sweeney D.J. and Williams T.A. (2008) *Statistics for Business and Economics*, 10th edn, Thomson Higher Education, Mason, OH.

Conover, W.J. (1999), *Practical Nonparametric Statistics*, 3rd edn, John Wiley & Sons, Inc., New York.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

product moment correlation

George R. Franke

The Pearson product moment correlation coefficient is an index of the degree of linear relationship between two variables. It is named after British statistician Karl Pearson, who was stimulated by Sir Francis Galton's conceptual and empirical work on heredity to develop the mathematics and use of correlations. For two variables X and Y, the correlation r is computed as

$$r = \frac{\sum (X - Mean[X])(Y - Mean[Y])}{\sqrt{\sum (X - Mean[X])^2 \sum (Y - Mean[Y])^2}} \tag{1}$$

The numerator in the equation reflects the tendency of X and Y to be related. If high or low X values are associated with high or low Y values, the products of these terms will be higher (for positive correlations) or lower (for negative correlations) than if the X and Y values were associated at random. The sum of these products depends on the measurement scales of X and Y, such as monetary values recorded in dollars versus euro versus yen. Therefore, the covariation in the numerator is expressed relative to the individual variation of X and Y in the denominator to make r scale free and constrained to be less than or equal to 1 in absolute value.

A correlation coefficient is equivalent to a standardized regression coefficient in a model with a single predictor variable (*see* MULTIPLE REGRESSION). Therefore, r can be interpreted as the predicted standard deviation difference in Y for two observations that are different by one standard deviation in X. The strength of the relationship can be interpreted in terms of r^2, the coefficient of determination, which gives the variance in one variable that is attributable to another variable. Rules of thumb for interpreting r are that population values of $r = 0.1$, 0.3, and 0.5 (positive or negative) can be viewed as small, medium, and large effects, respectively (Cohen, 1988). Estimated r values in samples from the population may be larger or smaller than the true value owing to sampling error. The significance of r in a sample of n observations can be tested by comparing

$$t = \sqrt{\frac{r^2(n-2)}{(1-r^2)}} \tag{2}$$

with values from Student's t distribution with n–2 degrees of freedom. In planning a study, using a sample size of $n = 8/r^2$ gives an 80–90% chance of obtaining significant results at $\alpha = 0.05$ for a given population value of r (Dunlap and Myers, 1997; *see also* STATISTICAL APPROACHES TO DETERMINING SAMPLE SIZES).

FACTORS AFFECTING r

In general, r can range from −1 to +1. However, if the distributions of X and Y are not equal, the maximum absolute value of r is less than 1. If the relationship is nonlinear, calculating r for transformed variables, such as the square root or logarithm of X or Y, may improve interpretability. Range restrictions, such as limiting a sample to observations with above-average scores on X, can dramatically reduce estimates of r. Dichotomizing a continuous variable, such as converting raw scores to 0 for below the median and 1 for above the median, can also substantially constrain estimates of r. Correlating ratios (e.g., comparing per capita measures between countries or states) can make correlations appear artificially high. Difference scores (e.g., between one time period and another) and measurement errors tend to reduce correlations.

Combining samples may inflate or deflate estimated correlations. For example, if r is positive for men but negative for women, a sample of both sexes may yield an estimate of $r = 0$. Or r could be negative for both men and women separately, but positive for both sexes together. Outliers – atypical observations – may also have a variety of effects, including changes in the magnitude, significance, and sign of estimated r values.

Thus, while r is a useful summary of linear relationships between two variables, it may conceal more than it reveals depending on the nature of the data. When raw data are available for analysis, creating a scatter plot of the X and Y values is a useful tool to

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

avoid misleading inferences from estimated correlation coefficients.

Bibliography

Cohen, J. (1988) *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn Lawrence Erlbaum Associates, Hillsdale.

Dunlap, W.P. and Myers, L. (1997) Approximating power of significance test with one degree of freedom. *Psychological Methods*, **2**, 186–191.

random coefficients modeling

Neeraj Arora

In marketing, it is important to recognize interpersonal differences in model parameters such as brand preference and price sensitivity. Unlike classical approaches in economics that view heterogeneity in model parameters as statistical nuisance, key marketing decisions such as product design and pricing mandate a careful understanding of interpersonal differences. In recent years, random coefficient modeling has significantly simplified our ability to characterize heterogeneity. A particularly useful aspect of random coefficient models that are discussed below is that each provides individual-level inference. This is critical for the field of marketing because not only are we concerned about the shape of the heterogeneity distribution but also how individual parameters vary with covariates such as demographics.

Consider the case of LOGIT MODEL, where the choice probability of alternative j for household i is $P_i(j) = \frac{e^{x_j'\beta_i}}{\sum_k e^{x_k'\beta_i}}$, where x_j is a vector of explanatory variables for alternative j and β_i is the vector of response coefficients for household i. To model unobserved heterogeneity, one can assume that individual response coefficients (β_i) come from an underlying distribution, sometimes referred to as the *mixing distribution* (*see* UNOBSERVED HETEROGENEITY; MODELS FOR CATEGORICAL DEPENDENT VARIABLES).

In the finite mixture approach (Kamakura and Russell, 1989), this distribution is discrete, that is, individuals are assumed to fall into M finite components (ϕ_m,β_m) such that $\sum_m \phi_m = 1$ and individuals belonging to component m have homogeneous preferences captured by the vector β_m. The number of components can be subjective and is often determined on the basis of a statistical fit criterion. An attractive aspect of this approach is that it nicely links to the theories of segmentation and has received considerable attention among academicians and practitioners in the 1990s (*see* LATENT CLASS AND FINITE MIXTURE MODELS). If the true underlying distribution of preference is continuous, the point mass approach of a finite mixture model can be restrictive. A simple approach to characterize heterogeneity in such a case may be to use a multivariate normal distribution (Allenby and Lenk, 1994). Because of its simplicity and conceptual appeal, the multivariate normal distribution of heterogeneity has experienced considerable popularity in marketing.

An even more flexible approach may be to use a normal component mixture model (Allenby, Arora, and Ginter, 1998). That is, $\beta_i \sim \sum_m \phi_m$ Normal $(\overline{\beta}_m, D_m)$, where m indicates the number of components and ϕ_m is the mass of each component. The normal component mixture model is capable of representing a wide variety of distributions of heterogeneity. A skewed distribution, for example, can be obtained with a two-component model with the mass point ϕ_2 small relative to ϕ_1, resulting in a thicker tail near the mean of the second component. As the diagonal elements of D_m approach zero, this model converges to a finite mixture model of heterogeneity.

In general, continuous distributions do a good job of characterizing the tails of heterogeneity distributions and therefore result in superior choice predictions (see Wedel *et al.*, 1999 for a review, and Andrews, Ainslie, and Currim, 2002). While the above discussion uses the multivariate normal as the kernel, other distributions such as gamma or beta may be better suited for applications, where the parameter space is restricted to be greater than zero or between 0 and 1, respectively. Another emerging approach to characterize heterogeneity uses the Dirichlet process mixture model prior (Ansari and Mela, 2003), which captures the uncertainty about the true functional form of the population distribution using a semiparametric approach.

To carry out estimation, finite mixture models use maximum likelihood methods. For continuous distributions of heterogeneity, Markov chain Monte Carlo (MCMC) estimation methods (Rossi, Allenby, and McCulloch, 2005) are commonly used. A particular advantage of MCMC methods is that they offer significant flexibility in specifying the functional form of the heterogeneity distribution. This is critical because accurate characterization of heterogeneity is an important consideration in order to develop models that are valid and actionable.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Bibliography

Allenby, G.M., Arora, N., and Ginter, J.L. (1998) On the heterogeneity of demand. *Journal of Marketing Research*, **35**, 384–389.

Allenby, G.M. and Lenk, P.J. (1994) Modeling household purchase behavior with logistic normal regression. *Journal of the American Statistical Association*, **89** (428), 1218–1231.

Andrews, R.L., Ainslie, A., and Currim, I.S. (2002) An empirical comparison of logit choice models with discrete versus continuous representations of heterogeneity. *Journal of Marketing Research*, **39**, 479–488.

Ansari, A. and Mela, C.F. (2003) E-customization. *Journal of Marketing Research*, **40** (2), 131–145.

Kamakura, W.A. and Russell, G.J. (1989) A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research*, **26**, 379–390.

Rossi, P.E., Allenby, G.M., and McCulloch, R. (2005) *Bayesian Statistics and Marketing*, John Wiley & Sons, Ltd, West Sussex.

Wedel, M., Kamakura, W., Arora, N. *et al.* (1999) Heterogeneity and Bayesian methods in choice modeling. *Marketing Letters*, **10** (3), 219–232.

repeated measures ANOVA

Mayukh Dass

In marketing research, one of the sources of variability in the observed measures is the respondents themselves. The respondents come from different backgrounds and have unique individual characteristics, and thus can perceive a study differently. Since this sample level variability reduces the effectiveness of the study, it is very important to separate it from the treatment effects and the experimental errors. Unfortunately, these differences among subjects are uncontrolled and are treated as errors in a between-subject design. Therefore, we use a within-subject design where the respondents are exposed to more than one treatment condition and hence repeated measures are obtained. There are also other instances when such repeated measures are necessary. For example, in a longitudinal study, respondents are measured repeatedly in different time periods. Take the case of a brand loyalty study, where customers are asked about their level of loyalty on a particular brand, once every 6 months. In this case, the dependent variable, that is, "brand loyalty" is measured repeatedly twice a year from the same subject. Repeated measures also play a vital role when it is difficult to recruit respondents. In such cases, we expose the respondents to different conditions and take repeated measures instead of recruiting new subjects.

Traditional analysis of variance or ANOVA (*see* ANALYSIS OF VARIANCE AND COVARIANCE to learn more about ANOVA) is not adequate to analyze this type of data as it is unable to control the correlation among the repeated measures. Therefore, a different analysis called *repeated measures ANOVA* is used. Repeated measures ANOVA is considered as an extension of a paired-sample t-test to a case of more than two related samples (Malhotra, 2007). To illustrate the underlying steps of the analysis, let us consider the case of single-factor repeated measures. Here, the total variation with $n(c-1)$ degrees of freedom is divided into a between-respondent variation and a within-respondent variation. The between-respondent variation with $(n-1)$ degrees of freedom captures the difference in means among the sample, and the within-respondent variation with $n(c-1)$ degrees of freedom captures the differences in the responses by the same respondent. The within-respondent variation is further divided into differences between the treatment means with $(c-1)$ degrees of freedom and the error variation with $(c-1)$ $(n-1)$ degrees of freedom.

$$SS_{\text{total}} = SS_{\text{between-respondent}} + SS_{\text{within-respondent}}$$

$$SS_{\text{within-respondent}} = SS_{\text{treatments}} + SS_{\text{error}}$$

Using the within-respondent variation, we can now use the following F-test to test for equal means

$$F = \frac{SS_{\text{treatment}}/(c-1)}{SS_{\text{error}}/(n-1)(c-1)} = \frac{MS_{\text{treatment}}}{MS_{\text{error}}}$$

where $SS_{\text{treatment}}$ = sum of squares of the treatment, SS_{error} = sum of squares of the error, $MS_{\text{treatment}}$ = mean sum of squares of the treatment, and MS_{error} = mean sum of squares of the error.

Bibliography

Malhotra, N.K. (2007) *Marketing Research: An Applied Orientation*, 5th edn, Pearson Education, Inc., Upper Saddle River.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

stepwise regression

George R. Franke

Stepwise regression is both a general term and a specific method for choosing predictor variables from a larger pool of possible predictors in MULTIPLE REGRESSION. Eliminating redundant or irrelevant variables increases parsimony and may reduce the variance of estimated regression coefficients and predicted values (*see* MULTICOLLINEARITY). However, eliminating relevant variables biases the estimates for the remaining variables. Step-wise approaches generally use the same data for variable selection and model estimation, which can produce results that are not generalizable to other samples. Findings on the basis of stepwise approaches must therefore be interpreted with extreme caution.

Methods for Variable Selection

Many approaches have been developed for selecting predictors in multiple regression (Miller, 2002). Three of the most common are forward selection, backward elimination, and (in its narrow sense) stepwise regression. In forward selection, the best predictor is chosen from all available predictors, followed by the second-best predictor given the first predictor chosen, and so on. Backward elimination starts with the entire set of predictors and removes the worst predictor, followed by the second-worst predictor given the other predictors still in the model, and so on. Stepwise regression modifies forward selection to test the relevance of variables already in the model after each new predictor is added, and omits predictors that no longer meet criteria for inclusion. Statistical packages for regression generally allow users to specify criteria for variable selection, such as significance levels or F-values that must be met to enter or remove predictors. Other criteria include R^2, the variance explained by the predictors; adjusted R^2, which lowers R^2 as a function of the sample size and number of predictors; and Mallow's C_p, which is a measure of the model's prediction error that approximately equals $p + 1$ for correct models with p predictor variables.

These approaches may identify different sets of predictors as optimum according to the criteria specified. None of the sets will necessarily be the best of all possible combinations of predictors. Some regression packages can identify the best possible one-variable model, two-variable model, and so on, according to the chosen criteria. The number of models to evaluate equals 2^p, so with very large values of p, this approach may not be feasible. However, with p of at least 50 or 60, modern computers can perform the analysis in minutes.

Problems and Recommendations

Because of random sampling error, the best set of predictors in one sample will often not be the best set in another sample or in the underlying population. Unless all p predictors are included in the analysis, significance tests and R^2 values may be far more liberal than their nominal values (McIntyre *et al.*, 1983). Some of the estimated coefficients are also likely to be substantially overstated relative to their population values.

For a given sample of observations, researchers using stepwise procedures may overlook how many other sets of predictors do almost as well as the selected set. Some of the alternatives are likely to perform better in other samples of observations. A simple check on the robustness of stepwise findings involves taking multiple random samples from the available data, and repeating the analysis on each subsample. Predictors that perform well in a majority of samples are good candidates for interpretation and application (Austin and Tu, 2004).

Stepwise regression procedures can be helpful tools for identifying important subsets of predictor variables. Researchers using them will generally get better results when they have large samples of observations and relatively small sets of potential predictors; when they consider multiple procedures and criteria for selecting predictors; when they cross-validate findings with new data or with random subsamples from the available data; and when they avoid self-deception in interpreting the findings.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Bibliography

Austin, P.C. and Tu, J.V. (2004) Bootstrap methods for developing predictive models. *The American Statistician*, **58**, 131–137.

McIntyre, S.H., Montgomery, D.B., Srinivasan, R., and Weitz, B.A. (1983) Evaluating the statistical significance of models developed by stepwise regression. *Journal of Marketing Research*, **20**, 1–11.

Miller, A. (2002) *Subset Selection in Regression*, 2nd edn, Chapman & Hall/CRC, New York.

logit model

Neeraj Arora

Consumer choice is arguably the most important variable in marketing. Examples include brand choice in a grocery store purchase, selection of a kitchen appliance at a retailer, or choosing a health club from the available alternatives. Choice is a discrete variable (taking a 0–1 value) and therefore does not lend itself to a linear regression model (*see* MULTIPLE REGRESSION; MODELS FOR CATEGORICAL DEPENDENT VARIABLES), where the dependent variable is continuous. On the basis of McFadden's (1974) random utility model of consumer choice, the utility for an alternative j in a choice set is

$$u_j = \mathrm{x}_j'\beta + \varepsilon_j \tag{1}$$

where x_j is a vector denoting the features of alternative j, β is a parameter vector reflecting the person's preference structure, and the errors ε_j are assumed to have independent Gumbel distribution. For this model, the probability of choosing an alternative j can be shown as

$$\mathrm{P}(j) = \frac{\mathrm{e}^{\mathrm{x}_j'\beta}}{\sum_k \mathrm{e}^{\mathrm{x}_k'\beta}} \tag{2}$$

In the brand-choice example, typical independent variables in a logit model include price, in-store promotions such as a special display and out-of-store promotions such as manufacturers coupon. The logit model is of great value because it allows a manager to understand the link between marketing actions and consumer behavior (Guadagni and Little, 1983). For example, the model can be used to conduct a "what if" analysis for a price change, an in-store display, or a coupon drop. Similar to the intercept term in a regression model, the logit model permits alternative or brand-specific dummies. The parameters corresponding to these dummies can be viewed as brand preference and provide a measure of relative preference of each brand. Some authors model state dependence in the logit model by including lagged terms (e.g., brand or flavor purchased in previous periods) as independent variables to capture interesting behavioral constructs such as brand loyalty, habit persistence, and variety seeking (Erdem, 1996; Seetharaman, 2004).

Two popular application areas for the logit model in marketing include household-level panel data (Fader and Hardie, 1996) and choice-based conjoint analysis (Green and Srinivasan, 1990). The former are based on data collected by Information Resources, Inc. (IRI) and The Nielsen Company for most packaged goods available from grocery stores, mass merchandisers, and drug stores. The latter entails data collected in a choice experiment by carefully manipulating features of product alternatives across a series of choice tasks (*see* CONJOINT ANALYSIS). For a logit model, a popular measure of how well the model fits the data is the hit rate. This measure is based on the simple calculation of the percentage of times that the model prediction of the chosen alternative is the same as the actual choice.

While enormously popular, the logit model has limitations. We note two such limitations. First, the distribution assumption for the error term is fairly ad hoc and is guided by the convenience of obtaining a closed-form expression in Equation (2). Unfortunately, this choice of error distribution also results in an undesirable property called the *IIA* or *independence of irrelevant alternatives property*. From Equation 2 it is easy to see that

$$\frac{\mathrm{P}(j)}{\mathrm{P}(k)} = \frac{\mathrm{e}^{\mathrm{x}_j'\beta}}{\mathrm{e}^{\mathrm{x}_k'\beta}} \tag{3}$$

Notice that the ratio of choice probabilities for two alternatives is independent of other alternatives in the choice set. So, an implication of the logit model is that alternative similarity does not affect choice probability. This is clearly an undesirable property for contexts that involve similar alternatives – such as a consumer selecting between Coke, Pepsi, RC Cola, and Sprite – where they may be more likely to switch from Pepsi to Coke than to Sprite. A probit model with normal and correlated errors could be used to overcome this limitation. Another option is to use a nested logit model (Ben-Akiva and Lerman, 1985; Train, 2003). A second limitation is the implicit assumption that consumers use a compensatory evaluation process (Equation 1 in

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

selecting an alternative), that is, they trade-off all attributes against each other. It is certainly plausible that consumers, in certain contexts, may use a noncompensatory evaluation process. A conjunctive model of consumer choice (Gilbride and Allenby, 2006) could be used to overcome this limitation.

Bibliography

Ben-Akiva, M. and Lerman, S. (1985) *Discrete Choice Analysis*. MIT Press, MA.

Erdem, T. (1996) A dynamic analysis of market structure based on panel data. *Marketing Science*, **15** (4), 359–378.

Fader, P.S. and Hardie, B.G.S. (1996) Modeling consumer choice among SKU's. *Journal of Marketing Research*, **33**, 442–452.

Gilbride, T.J. and Allenby, G.M. (2006) Estimating heterogeneous EBA and economic screening rule choice models. *Marketing Science*, **25** (5), 494–509.

Green, P.E. and Srinivasan, V. (1990) Conjoint analyses in marketing research: new developments and directions. *Journal of Marketing*, **54**, 3–19.

Guadagni, P.M. and Little, J.D.C. (1983) A logit model of brand choice calibrated on scanner data. *Marketing Science*, **2**, 203–238.

McFadden, D. (1974) Conditional logit analysis of qualitative choice behaviour, in *Frontiers in Econometrics* (P. Zarembka), Academic Press, New York.

Seetharaman, P.B. (2004) Modeling multiple sources of state dependence in random utility models: a distributed lag approach. *Marketing Science*, **23** (2), 263–271.

Train, K.E. (2003) *Discrete Choice Methods with Simulation*, Campbridge Press.

unobserved heterogeneity

Neeraj Arora

Heterogeneity, or interpersonal differences between individuals, could be categorized into two groups: observed and unobserved. Observables such as consumer demographics (age, gender, geographic location) could be included in a model to establish the link between model parameters and observed data. This is often referred to as *observed heterogeneity*. The remaining heterogeneity in model parameters, unaccounted for by measures available to the analyst, is called *unobserved heterogeneity*. In marketing, it is important to recognize and model unobserved heterogeneity when possible.

In general terms, a model could be viewed as $y = f(x, \beta, \varepsilon)$, where we attempt to explain the dependent variable y as a function of independent variables x and model parameters β (*see* MULTIPLE REGRESSION). Unexplained variation in y is captured via ε, the error term. For example, in a logit model, the dependent variable is choice and the errors are assumed to follow an independent Gumbel distribution (*see* LOGIT MODEL). For a given person i, the choice probability of choosing alternative j could then be written as $P_i(j) = \frac{e^{x_j'\beta_i}}{\sum_k e^{x_k'\beta_i}}$. In its simplest form, one could fit an aggregate model of consumer choice (Ben–Akiva and Lerman, 1985; Guadagni and Little, 1983) by assuming a response function that is homogeneous across the population of interest. The consumer choice in this model is explained entirely by the marketing mix elements included. Importantly, it is assumed that $\beta_i = \beta$ for all i.

A limitation of an aggregate model is that no attempt is made to assess how the response parameters vary across the population of interest. This highly restrictive assumption of constant model parameters across all units of analysis is undesirable for most marketing applications. For example, if the independent variables (x_j) in this model include price, in-store promotions (e.g., a special display), and out-of-store promotions (e.g., manufacturer coupons) then it is safe to assume that the corresponding model parameters (β) should be different for each person i. It is well known, for example, that consumers exhibit differential sensitivity to price changes and promotions.

When testing theory, economists often attempt to characterize the "average" consumer. Observed heterogeneity caused by variables z_i is typically handled by expressing $\beta_i = z_i\gamma$ and estimating γ instead of β. However, unobserved heterogeneity is often treated as a nuisance parameter. In contrast, careful characterization of unobserved heterogeneity is at the heart of most marketing decisions. For example, we often attempt to segment the market on the basis of interpersonal differences in price sensitivity, product feature (e.g., low fat, high fuel efficiency) sensitivity, and brand preference. Indeed product-line length decisions are inextricably linked to accurate characterization of unobserved heterogeneity. From a statistical standpoint, it is important to carefully model unobserved heterogeneity because aggregate models suffer from biased parameters (Blattberg and Neslin, 1990, p. 181) and subsequent marketing actions on the basis of biased estimates are likely to be suboptimal.

Ability to accurately characterize unobserved heterogeneity requires multiple observations for each unit of analysis. This is certainly plausible for syndicated data available from Information Resources, Inc. (IRI) and The Nielsen Company for most packaged goods available via grocery stores, mass merchandisers, and drug stores. Unobserved heterogeneity among purchasing households or stores, typically explains significant variation in any type of a model. Another context is conjoint analysis (ratings and choice based), where the very purpose of the experiment is to uncover the heterogeneity in relative importance of product features across the population of interest (*see* MULTIPLE REGRESSION). Recent innovations in computing power and Bayesian simulation-based methods of model estimation (Train, 2003; Rossi, Allenby, and McCulloch, 2005) have significantly simplified our ability to estimate models with unobserved heterogeneity (*see* RANDOM COEFFICIENTS MODELING).

Bibliography

Ben-Akiva, M. and Lerman, S. (1985) *Discrete Choice Analysis*, MIT Press, MA.

Blattberg, R.C. and Neslin, S.A. (1990) *Sales Promotion: Concepts, Methods and Strategies*, Prentice Hall.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Guadagni, P.M. and Little, J.D.C. (1983) A logit model of brand choice calibrated on scanner data. *Marketing Science*, **2**, 203–238.

Rossi, P.E., Allenby, G.M., and McCulloch, R. (2005) *Bayesian Statistics and Marketing*, John Wiley & Sons, Ltd, West Sussex.

Train, K.E. (2003) *Discrete Choice Methods with Simulation*, Cambridge Press.

customer-satisfaction research

Vikas Mittal

Introduction

Customer-satisfaction research is widely undertaken by marketing-research departments worldwide. The logic for conducting customer-satisfaction research is undeniable. Customers are the biggest source of cash flow for a company. If a majority of a firm's customers are satisfied, then a variety of positive outcomes occur. Satisfied customers are likely to continue repurchasing a firm's offerings, purchase more from the firm, engage in more cross-buying, and have lower service and retention costs. Satisfied customers help a firm to lower the cost of customer acquisition through positive word of mouth and recommendations to friends and family. They also have lower price elasticity, that is, are less likely to defect when competitors offer lower prices. Finally, they are also more forgiving: when there is an occasional product or service failure; highly satisfied customers may attribute it to external causes and stay loyal to the firm. New research also shows that a strong customer-satisfaction reputation reinforces a firm's reputation in areas such as corporate-social responsibility and financial performance, leading to a virtuous cycle of positive performance. No wonder firms invest heavily in customer-satisfaction research to monitor and assess the satisfaction level of their customer base.

Customer-satisfaction studies in firms, typically, are conducted in the form of cross-sectional surveys. Results from these surveys can be aggregated over time as tracking studies. In addition to this research done by firms, a sizable amount of customer-satisfaction research has been conducted at the macro level as well. This macro-level research has provided marketing managers with a credible and robust view of the satisfaction–financial performance linkage and is described next.

Customer-satisfaction Research: Strategic Overview

Macrolevel customer-satisfaction research that uses a firm as the unit of analysis is primarily available through the American Customer-Satisfaction Index (ACSI) at the University of Michigan Business School. The ACSI was developed by the National Quality Research Center at the University of Michigan to obtain a customer-based measure of firms' ability to satisfy customers (Anderson, Fornell and Mazvancheryl, 2004). A representative sample of approximately 250 current customers from each firm is interviewed each year using computer-assisted telephone interviews. A different sample of customers is contacted each year. All respondents have purchased and used products or services from the firm in a defined period of time. For a given year, the database contains over 200 000 customer surveys. Twenty-five industries from the retail, finance/real estate/insurance, durable manufacturing, nondurable manufacturing, basic services, and transportation/communications/utilities sectors are included in the ACSI dataset. The following industries are represented in the sample (a subset of the firms in the sample is noted in parentheses): apparel (Liz Claiborne, Levi Strauss); athletic shoes (Nike, Reebok); automobiles (Chrysler, Honda); banks (Wells Fargo, Key Bank); brewing (Anheuser-Busch); department stores (Federated, May); discount stores (K-Mart, WalMart); fast food (McDonald's, Wendy's); food processing (Dole Foods, Heinz, Kellogg, Quaker Oats); hotels (Promus); household appliances (Maytag, Whirlpool); soft drinks (Coca-Cola, Pepsico); personal care (Dial, Colgate-Palmolive); personal computing (Compaq; Dell); pet foods (Ralston); service stations (Mobil, Texaco); supermarkets (Kroger, Safeway); tobacco (Phillip Morris); telecommunications-long distance (MCI, Sprint); telecommunications-local (BellSouth, US West).

In the ACSI measurement system, customer-satisfaction ratings are obtained from customers of each individual firm and used to create a customer-satisfaction index that can range from 0 to 100, with 100 as the highest level of satisfaction. A key advantage of this measure is the methodological consistency across all firms and over time. In other words, exactly the same survey instrument, interviewing methodology,

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

and statistical techniques are applied. An individual firm can thus benchmark its performance against competitors in the same industry as well as monitor performance relative to other industries.

Several academic studies have established a strong association between customer satisfaction and firm financial performance measured as long-term shareholder value, return on investment (ROI), return on assets (ROA), cash flow, risk, bondholder return, and so on. These studies also document that the strength of association can vary by many factors such as industry, firm efficiency, and diversification. One study examining the association between firm-level customer satisfaction and long-term shareholder value found a strong association in industries such as department stores, supermarkets, appliances, life insurance, consumer electronics, but relatively weak association in industries such as property insurance, food processing, personal computers, apparel, and automobile. Scholars have also shown that customer-satisfaction scores at the firm-level are best understood within the context of other factors such as the firm's corporate-social responsibility behavior as well as diversification strategy.

This macrolevel research using the ACSI is particularly important for marketing researchers as they can use it to establish the initial connection between a firm's satisfaction scores and financial performance. This is relatively easy to do, as firm-level satisfaction scores can be downloaded from the ACSI website. If needed, industry-level trends can also be monitored using the ACSI. Building on this initial evidence, marketing departments can make the case for their company to conduct customer-level research to measure and monitor satisfaction.

Conducting Customer-satisfaction Studies for a Firm

At the heart of every satisfaction research is a carefully designed satisfaction survey. Particular attention should be paid to the survey content and design, design of the sample, survey methodology, and the analysis plan.

Satisfaction survey design and content. Though the satisfaction survey for each firm must be adapted to its specific situation, a typical satisfaction survey consists of four parts. These are (i) measures of overall satisfaction, (ii) measures of behavioral intentions, (iii) attribute-performance perceptions, and (iv) background information.

Overall customer satisfaction. Professor Richard L. Oliver did pioneering work in defining satisfaction from the customer's perspective. Customer satisfaction is a summary judgment that has both an attitudinal and affective component, and which describes the cognitive and affective reaction of customers toward their consumption experience. In a typical customer-satisfaction survey, the construct is measured by asking customers to rate their overall experience with the company's product or service. Firms use a variety of overall satisfaction items such as

a. Overall, how satisfied are you with . . .' (1 = extremely dissatisfied, 5 = extremely satisfied).
b. Please rate your agreement with the following item: "I am very satisfied with . . ." (1 = strongly disagree, 10 = strongly agree).
c. How would you rate your experience with . . .' (1 = poor, 7 = excellent).

There are dozens if not scores of different overall satisfaction scales, each with its own advantages and disadvantages. Empirical research, however, shows that different scales tend to perform adequately and should be chosen on the basis of the context of the particular firm. More generally, a firm should pretest a few different scales to assess their suitability for its purpose. From a practical standpoint, it is important to ensure that the scale has face validity and finds acceptance within the firm especially the marketing department and frontline employees who would most likely use the results of the customer-satisfaction study.

Behavioral intentions. These include measures such as likelihood to repurchase, likelihood to recommend, and likelihood to complain. Similar to customer satisfaction,

behavioral intentions can be worded in many different ways:

a. How likely are you to repurchase brand XXX in the next six months (1 = not at all likely, 10 = completely likely)?
b. Would you be likely to repurchase . . . (1 = definitely not, 5 = most definitely)?
c. Would you recommend this brand to your friends and family (1 = definitely will not recommend, 7 = will definitely recommend)?

Sometimes, firms also measure retrospective self-reports of behavior. These can provide useful insights as well. For instance, research has shown that overall satisfaction ratings predicted self-reported word-of-mouth activity among a national sample of customers. Similarly, researchers have also shown customer satisfaction to predict self-reported share of wallet (SOW).

Attribute-level perceptions. Typically, the survey measures customer evaluations on various attributes. For example, a customer-satisfaction survey for the automotive industry may measure performance on attributes such as "interior roominess," "fit and finish," "reliability," and so forth. A banking survey may include attributes like "waiting-time in line," "respect by employees," and so forth. A satisfaction survey for health care may include "correct diagnosis," "doctor doesn't rush me," and "doctor treats me with respect." Performance ratings on each attribute can be obtained using a variety of scales:

a. Rate the performance of each attribute as "excellent = 5, very good = 4, good = 3, fair = 2, poor = 1."
b. How did the product perform based on your expectations? Responses are obtained as: (above my expectations, met my expectations, below my expectations). This is a typical expectation disconfirmation scale, though, used by many firms to measure attribute performance.

An important aspect of developing the attribute list is to ensure that the list incorporates the customer's point of view. For this, marketing researchers should ensure the following:

- The items should not just use the manager's judgment. Instead, qualitative research such as focus groups should be used to surface key attributes of the product and/or service from the customer's point of view.
- The items generated should not overlap conceptually. For instance, a car dealership may measure "politeness" and "courteousness" of the service staff. However, if these two are highly correlated, it may pose problems in the multivariate analysis because of multicollinearity.
- The items should be worded using terms that are familiar to, and easily understood by the customers. Thus, technical jargon should be avoided in wording the attributes.

Customer background variables. In addition to demographics like gender, age, income and so forth, a satisfaction survey may also measure ownership of competitive brands (in an automotive study), or balance at other banks. These latter measures enable a firm to engage in competitive assessment. Typically, these background variables can be used to segment the customer base for subcell analysis.

Survey sample. A typical sample for a customer-satisfaction study includes customers who have purchased and/or used the firm's offering in the past, during a prespecified period of time. Such an approach lets firms get a snapshot of their current customers. However, it is also useful to measure customer-satisfaction perceptions among past customers, that is, customers who are no longer with the firm. Similarly, measuring customer satisfaction among potential customers is also useful because they are the customers who typically use competitive offerings. Thus, in an ideal situation, the firm would conduct a satisfaction survey with the following samples:

- Current customers who are using the firm's offering.
- Past customers who no longer use the firm's offerings. They are also termed as "*lapsed customers*," "*former customers*," or "*defectors*."

- Potential customers who currently use competitive offerings but not the focal firm's brand. In many cases, potential customers could have been past customers of the firm.
- Finally, for our services (e.g., banking and insurance) and products (e.g., automobile) is it possible for customers to simultaneously use multiple brands? In such cases, firms should measure overall satisfaction and behavioral intentions toward all the relevant brands.

Broadening the sample in this way is strategically useful in many ways. A comparison of customer-satisfaction scores and intentions between current and potential customers provides competitive benchmarking for overall satisfaction score, and attribute-performance scores. Comparing current and past customers can provide insights into why customers may have left the firm. Some studies have also found differences in customer satisfaction based on how long a current customer has been with a firm. Expectedly, customers who have been with a firm longer are more satisfied than customers who have been with the firm for shorter durations. Further, the relative importance of attributes varies on the basis of tenure. In the automotive industry, product performance is more important early on, but dealership importance becomes more important after three years of ownership.

Studies have also shown some differences in customer-satisfaction scores based on demographics, though the nature and direction varies by industry. Thus for many durable goods, females exhibit higher scores than males, older customers are more satisfied than younger customers, and higher income/education shows a negative association with customer-satisfaction scores. Thus in analyzing customer-satisfaction scores, it is useful to statistically control for these factors.

Survey methodology. A satisfaction survey can be conducted in many ways, with mail, telephone, and Internet being the most prominent methodologies (*see* SURVEY RESEARCH). Clearly, each methodology has its own advantages and disadvantages. Compared to a telephone survey, a mail survey is cheaper. However, it is less versatile than a phone or Internet survey in that skip patterns and item randomization are not possible in a mail survey. However, compared to a phone survey, a mail survey affords the respondent privacy and reduces demand bias due to acquiescence. Some studies have found that satisfaction scores from telephone surveys tend to be higher than mail surveys because respondents are less likely to say negative things about the company to another person. A telephone survey, while being more expensive, is faster than a mail survey, and affords the interviewer opportunities to probe and follow up the responses provided. For instance, the interviewer can probe customers who indicate they are very dissatisfied. This flexibility is unavailable in a regular mail survey. Finally, Internet or web surveys are becoming popular as a relatively inexpensive way to reach a mass audience. Especially for firms having a robust database of email addresses, this can be an attractive survey methodology that combines the virtues of a mail survey with the speed of a telephone survey.

Overall, each methodology has its strengths and weaknesses. When cost is not a concern, a telephone methodology would be recommended. However, mail and online surveys by virtue of nonintrusiveness may provide more accurate data. In understanding these issues, especially cost, it is also important to consider the unintended positive benefits of such research. Marketing scholars have done a series of studies showing that customers filling out a satisfaction survey – irrespective of the rating they give – not only buy more from the firm later on, but are also more profitable. The basic logic is that being requested by the firm to be part of a customer-satisfaction survey sends a signal to the customers that the firm cares. This, in turn, prompts positive behavior among customers. In fact, calculations show that virtually in all cases customer-satisfaction studies pay for themselves from increased revenue and profitability.

Analysis of customer-satisfaction surveys. The analysis of a customer-satisfaction survey can take many forms ranging from basic univariate analysis to advanced multivariate analysis. We can look at it in increasing order of complexity.

Univariate descriptive (cross-sectional). This is the most commonly used approach. Here either the average score or some variation of the "top box score" (e.g., top-2 or top-3 box) is used to describe the data (*see* PRODUCT MOMENT CORRELATION). Thus the average for top-2 box score is displayed as bar charts for overall satisfaction, attribute performance, and various behavioral intentions. Frequently, these are broken down by customer characteristics deemed relevant (e.g., by gender, age, brand ownership, number of years with the firm, and so forth).

Univariate descriptive (over time). If a firm conducts the customer-satisfaction survey over time (e.g., every year), the average scores or top box scores can be displayed over time to examine trends in the firms' scores on customer satisfaction, behavioral intentions, and attribute performance. Most firms use the tracking studies extensively to monitor their customer base over time.

Bivariate analysis. To understand the association between overall customer satisfaction and intentions, some firms complete the correlation between the two (*see* PRODUCT MOMENT CORRELATION). In some cases, the association of attribute-performance ratings with overall customer satisfaction is calculated using bivariate correlation. This is used to rank order attributes such that the attribute with the highest correlation with overall satisfaction is deemed most important. Statistically, this approach is quite problematic as it does not account for the underlying correlation structure in attribute ratings. Multivariate analysis is required for addressing multicollinearity.

Multivariate analysis. This is used most often to ascertain attribute importance (*see* MULTIPLE REGRESSION; MULTICOLLINEARITY). With overall satisfaction as the dependent variable, and attribute-performance ratings as the independent variable, a multiple regression model can be estimated. The coefficient of each attribute variable denotes its relative importance. The idea is that the higher the relative impact of an attribute on customer satisfaction, the more important it is as a driver of customer satisfaction. Practically speaking, such statistical models must account for the problem of multicollinearity since attribute-performance ratings obtained from the same survey tend to be highly correlated, sometimes owing to the halo effect.

Importance–performance analysis. This analysis uses the importance estimate from the multivariate regression, combined with the performance ratings on attributes (e.g., top-2 box score or average rating) to classify attributes into four groups:

- attributes that have high importance and on which the firm is performing high represent strengths of the firm;
- attributes that have high importance but on which the firm has low performance represent weaknesses that should be immediately addressed;
- attributes that have low importance but on which the firm has very high performance indicate that the firm may be overinvesting in that attribute. A strategy that may be adopted would be to increase the perceived importance of the attribute.
- attributes that have low importance and on which the firm has low performance, indicate that continuous monitoring is warranted lest either changes dramatically.

An importance–performance analysis results in a 2×2 map where the Y-axis shows the relative performance and X-axis shows the relative importance of each attribute. While different terms like "quadrant analysis," "strategic attribute mapping," "product-aspect maps" are used, the basic logic remains the same. Sometimes, on the same map it is useful to also plot the relative performance of "best-in-class" competitor for benchmarking purposes. Firms can use these importance–performance charts to develop concrete action plans for driving up customer satisfaction.

Smart firms take a return on quality (ROQ) approach to importance–performance analysis. Popularized by Professor Roland Rust and his colleagues, the ROQ approach suggests that rather than being an afterthought, the financial impact of all quality-related investments in customer satisfaction should be quantified.

They conducted a study at Chase Bank to show how investments in quality improvements, when measured, had a statistically significant impact on overall financial performance.

Using the Results of a Customer-satisfaction Study

How do firms use the results of a customer-satisfaction study? A recent, in-depth study of over 120 managers from 35 businesses provides critical insights in this regard. Among others, the study arrived at the following conclusions:

- Over 90% of the firms used their customer-satisfaction study to examine current attribute level and overall satisfaction scores compared to past scores. However, less than 20% examined overall satisfaction and behavioral intentions, and only 3% related customer satisfaction to actual customer behaviors.
- With respect to analysis, 92% of the firms conducted univariate analysis like describing the mean and top-2 box scores. Only 38% conducted any bivariate analysis such as two-way cross tabulations or correlation analysis. Surprisingly, less than 1 in 10 (8%) did any multivariate analysis.
- Customer-satisfaction surveys were used in evaluating customer service, but few formal linkages existed to integrate customer-satisfaction scores with employee evaluation, compensation/rewards, and product/service development.
- Only 40% of the firms used their customer-satisfaction survey as a tool to learn about customers. A major impediment to wider usage was lack of formal structures to disseminate the results of the customer-satisfaction survey within the firm. The study found that 24% disseminated the results quarterly and 27% disseminated them on a monthly basis. In 86% of the studies the customer-satisfaction reports and dissemination efforts were targeted to senior management. In 62% of the firms, frontline employees were the target of dissemination efforts.
- Frequently, potential users viewed the customer-satisfaction information as being provided "too late" for it to be usable. Even if provided on time, it was not seen as directly relevant to fixing customer-satisfaction issues from a customer's perspective.

The study also identified several discouraging practices, including but not limited to the following:

- The focus was only on current customers to the exclusion of past or potential customers.
- More than half the firms did not conduct key-driver analysis to identify important attributes, and only two firms in the study linked overall satisfaction to postpurchase intentions.
- While customer-satisfaction data are disseminated to employees, the employees lack the training to use the data to identify root causes or fix customer-satisfaction problems. This makes them skeptical of customer-satisfaction studies.
- Satisfaction data is used mostly at a tactical, rather than strategic level.

Given these findings, firms interested in strengthening their customer-satisfaction research need to not only look at the actual research methodology (e.g., survey design, analysis plan, and so on) but also examine the organizational structures and processes by which customer-satisfaction research is disseminated and utilized within the firm. Clearly, the latter task requires intervention from and the support of top management.

Emerging Issues in Customer-satisfaction Research

Over time, and more recently, many issues have emerged in the conduct of customer-satisfaction research. Some important issues are described below.

Customer satisfaction and net promoter score. Popularized by Reichheld as an alternative to measuring customer satisfaction, the Net Promoter score is based on a customer's likelihood to recommend the firm's offering to others. It is based on the difference between "promoters" (e.g., those giving a 4/5 on

a 5-point likelihood to recommend) and "detractors" (i.e., those giving 1/2). Many carefully conducted studies comparing the predictive ability of customer satisfaction and Net Promoter show the superiority of customer satisfaction in predicting key financial metrics. One large-scale study using the ACSI conclusively showed that customer-satisfaction scores predicted cash flow, long-term shareholder value, and other stock market metrics, while the Net Promoter score had zero predictive ability. Scholars have also highlighted the theoretical inadequacy and possible statistical problems in Reichheld's original analysis. Given this robust evidence, the use of Net Promoter score, in lieu of customer satisfaction, should be discouraged. It seems that Net Promoter score, as a concept, is not only conceptually and empirically flawed, but also that its use discourages firms from focusing on more useful concepts such as customer satisfaction.

Behavioral consequences of customer satisfaction. It is important to go beyond simple intentions that are concurrently measured in the same survey as customer satisfaction for two reasons. First, the halo effect can inflate the observed association between customer satisfaction and intentions. Secondly, intentions and behavior are not the same construct. Thus, imputing the behavioral consequences of customer satisfaction based on intentions may be misleading. The current thinking in customer-satisfaction research is that firms should collect actual behavioral metrics from internal records and link them to customer satisfaction. For instance, using actual behavioral data (and not just self-reports), studies have found a link between customer satisfaction and (i) repurchase of automotive brand, (ii) bank balance, and (iii) customer tenure at many firms. Recently, scholars have shown that customer satisfaction is not only linked to market share, but more importantly to the SOW, especially in financial services. More importantly, the functional form linking customer satisfaction to intentions (e.g., decreasing returns) can be very different than the functional form linking customer satisfaction to actual behavior (e.g., increasing returns). Hence, when available, one should use actual behavior and not intentions as the dependent variable.

The satisfaction–profit chain. In an attempt to apply customer-satisfaction research more comprehensively and strategically, the satisfaction–profit chain (SPC) is used as an underlying template. A typical satisfaction survey links attribute performance to overall satisfaction and overall satisfaction to behavioral intentions. The SPC links attribute-performance perceptions to operations inputs. Thus, "waiting-time in line" may be linked to the number of tellers per branch. Attribute performance is linked to overall satisfaction, which is linked to behavioral intentions. Downstream, behavioral intentions are linked to actual customer behaviors and, finally, financial metrics like sales and profitability. Such a comprehensive modeling approach is being increasingly used in financial services, hospitality, healthcare, and business-to-business services. For instance, useful statements like the following can be made: A $2 million investment in teller training is linked to 0.25 unit change in "tellers are polite." Increasing customer satisfaction by 0.5 units translates into $3.8 in increased bank balance per customer and $0.04 change in profit per customer. Using such information, the firm can understand the full impact of its customer-satisfaction investments. This is fully consistent with the "return on customer satisfaction" approach which argues that such spending is best viewed as a strategic investment and not an expense.

Customer satisfaction and service quality. Recognizing that service quality and customer satisfaction are conceptually different constructs is important. The initial debate about causal sequence has been resolved to a viewpoint that service quality is a key antecedent of customer satisfaction. Empirical studies also show both of them to be highly correlated with studies finding correlations in the range of 0.80–0.95. Thus while they should not be used interchangeably, there may not be a pressing need to separately measure them and expect different results from a practical point of view.

Customer satisfaction and employee satisfaction. There is an increasing trend to try and link a firm's employee-satisfaction scores to its customer-satisfaction scores. Despite the

widespread belief that employee satisfaction leads to firm profitability, studies have failed to find a direct link between the two. A recent study evaluated all the empirical evidence linking employee satisfaction to organizational financial performance. Except for two studies, all other studies found a zero correlation between employee satisfaction and financial performance measures like sales and profitability. Unfortunately, the belief that these two are related persists.

Instead, the study found that employee satisfaction is strongly related to customer satisfaction, and customer satisfaction robustly affects a firm's financial performance metrics. In other words, the effect of employee satisfaction on financial performance was fully mediated through customer satisfaction. This finding has profound implications for customer-satisfaction research. At the very least, it implies that both customer-satisfaction studies and employee-satisfaction studies should be analyzed together to fully understand the factors that drive a firm's profitability. In a study of Brazilian banks, the authors found human factors (e.g., satisfaction with tellers and bank managers) to be more consequential for profitability than technology factors (e.g., number of automatic teller machines at a bank branch). However, they also found that the cost of improving performance on human factors was 4 times as high as the cost of improving performance on technology factors. Clearly, when integrating employee satisfaction with customer satisfaction, and evaluating it as a causal factor, cost-related issues should also be considered.

Nonlinearities in satisfaction research. When estimating the relationships among constructs such as attribute performance, overall satisfaction, repurchase intentions, and repurchase behaviors, most marketing researchers assume their relationship to be linear. Recent research shows this may not be the case. Many of the relations are asymmetric, that is, a decrease from the midpoint has a stronger negative effect than the positive effect of a corresponding increase in attribute performance. For instance, a study of automotive customers found that a decrease in performance had a stronger impact on overall satisfaction than a corresponding increase in performance on attributes such as brakes (0.24 versus 0.00), transmission (0.35 versus 0.04), power and pickup (0.21 versus 0.06), and quietness (0.23 versus 0.11). Similarly, for health care, there was a stronger impact on satisfaction when performance was negative than when it was positive on attributes such as thorough/attentive physician (0.66 versus 0.11), spends time with me (0.40 versus 0.17), gives referrals (0.77 versus 0.07), and convenient office hours (0.37 versus 0.07). Clearly, on many such attributes, firms should strive to eliminate the negative before they try to accentuate the positive.

Research has also shown that many of the relationships particularly between overall satisfaction and its consequences are nonlinear, with some relationships showing increasing returns and others showing decreasing returns. Thus, a study of automotive customers showed that the link between satisfaction and repurchase intentions showed decreasing returns, while the link between satisfaction and repurchase behavior showed increasing returns. Similar nonlinearities have been found in industries such as banking, insurance, utilities, and healthcare. These industry-specific patterns should be incorporated in the analysis plan for a customer-satisfaction study.

Caution in Using Customer-satisfaction Research

Despite its many virtues, customer-satisfaction research should be used cautiously. Listed here are several issues about which firms should be careful:

- When ascertaining attribute importance, that is, the impact of an attribute's performance on overall satisfaction, many analysts use bivariate correlation analysis. This can be grossly misleading, and multiple regression analysis should be used whenever possible.
- Given budget constraints, firms typically survey their own customer base to measure satisfaction. By surveying a smaller sample of lost customers (i.e., customers no longer with the firm) and competitor customers, insights from satisfaction surveys can be substantially improved.

- Many customer-oriented firms use customer-satisfaction studies to set action priorities. For instance, realizing lower performance scores on "comfort of waiting area," a small regional bank's marketing team wanted to refurnish the waiting areas in all its branches. However, no consideration was given to the cost-benefit calculation. In other words, customer-satisfaction surveys do not take into account the investments needed to implement the findings. For this, an ROQ or a return on satisfaction (ROS) is recommended.
- The above is consistent with the SPC that integrates customer satisfaction with both operations and finance within a firm. In doing so, attention should be paid to the profit implications (not just revenue enhancements) that result from investments aimed at improving overall satisfaction.
- Some researchers use the term *loyalty* to refer to intention items, especially intent to recommend. However, careful studies of the concept of loyalty show that loyalty is a multidimensional term with no agreed-upon way to conceptualize and measure it. For instance, some scholars differentiate behavioral (actual behavior) and attitudinal loyalty (based on intentions and attitudes). Given this disagreement, we suggest that marketing studies refrain from using the term *loyalty*. Rather specific terminology such as "behavioral intentions" or "repurchase intentions", or "actual repurchase behavior" which is clear, specific and usable should be encouraged. In many cases, it is not known how repurchase intentions correlate with actual behavior. A large-scale study showed virtually no correlation between satisfaction and actual repurchase *behavior* for car buyers, even though satisfaction and repurchase *intent* were highly correlated.

Concluding Comments

Customer-satisfaction research has matured in its own right among market-oriented firms. As evidence mounts for the strong association between customer satisfaction and firm financial performance, firms need to move from using customer-satisfaction research as a descriptive tactic, to a strategic tool. In this article, we have described how a basic customer-satisfaction study is designed and conducted. More importantly, we have provided state-of-the art insights into its usage among firms incorporating the latest developments in academic research. We hope that this will help firms to not only evaluate their current practices, but also improve them in the near future.

Bibliography

Anderson, E.W., Fornell, C., and Mazvancheryl, S. (2004) Customer satisfaction and shareholder value. *Journal of Marketing*, **68** (4), 172–185.

Anderson, E.W. and Mittal, V. (2000) Strengthening the satisfaction-profit chain. *Journal of Service Research*, **3** (2), 107–120.

Bolton, R.N. (1998) A dynamic model of the duration of the customer's relationship with a continuous service provider: the role of satisfaction. *Marketing Science*, **17** (1), 45–65.

Kamakura, W.A., Mittal, V., de Rosa, F., and Mazzon, J.A. (2002) Assessing the service-profit chain. *Marketing Science*, **21** (3), 294–317.

Morgan, N.M. and Leotte Rego, L. (2006) The value of different customer satisfaction and loyalty metrics in predicting business performance. *Marketing Science*, **25** (5), 426–439.

Oliver, R.L. (1997) *Satisfaction: A Behavioral Perspective on the Consumer*, McGraw Hill, Boston.

Rust, R.T., Zahorik, A.J., and Keiningham, T.L. (1995) Return on quality (ROQ): making service quality financially accountable. *Journal of Marketing*, **59**, 58–70.

consumer materialism

Marsha L. Richins

Consumer materialism refers to the importance that a consumer places on the acquisition and possession of material objects. In the consumer-behavior literature, materialism is usually considered to be a personal value, meaning that it affects the priorities in people's lives and the choices they make. Materialists believe that acquisition of material goods is a desirable and effective way to achieve important life goals. Thus, a materialist is more likely than others to see acquisition as a way to attain happiness, develop relationships with others, achieve status and a sense of self worth, and reach other important life goals. Richins and Dawson (1992) identified three elements that constitute materialism: the belief that acquisition is necessary for happiness, the tendency to judge the success of one's self and others by their possessions, and the centrality of acquisition and possessions in one's life. Because materialism is a value orientation and not a behavior, it is independent of financial status and is readily observed in both developed and developing economies. Materialism is most commonly measured by the material values scale (Richins, 2004).

Materialism should not be confused with conspicuous consumption, which is the ownership and display of status objects to enhance one's relative standing and inspire envy. While some materialists engage in conspicuous consumption, it is possible to have materialistic values without engaging in this particular behavior, particularly if status is not an important life goal for a consumer.

Although the formulation of materialism described above is the dominant one in consumer behavior, there are some alternative conceptualizations that should be recognized. Kasser (2002), a psychologist, has equated materialism with the pursuit of wealth, attractiveness, and social recognition and has looked especially at the negative effects of these goal pursuits on well-being. Inglehart (2008), a political scientist, has conducted extensive cross-national research over several decades to investigate the relative emphasis members of a society place on *material values* (defined as the importance one places on economic and physical security) and *postmaterial values* (the priority assigned to such things as free speech and a greater say in government decisions). A third approach (Belk, 1985) views materialism as a combination of the personality traits of envy, nongenerosity, and possessiveness.

Materialism has important implications for society as a driver of personal consumption, and thus of economic growth. It also has personal implications because of its negative association with well-being (Christopher, Saliba, and Deadmarsh, 2009), strength of personal relationships, and altruistic behavior. Because of materialism's importance, much has been written about its potential causes. Commonly, high materialism levels have been attributed to insufficiently bridled capitalism coupled with a consumer society that is endorsed and facilitated by marketing firms, assisted by mass media. A lack of spirituality is also often alluded to. The preponderance of writings on the causes of materialism is speculative in content, and in any event the construct itself and the systems in which it is embedded are too complex to assign a single causal variable. However, some empirical research has provided insight by examining potential influences on individuals' materialism levels, including media exposure (Shrum, Burroughs, and Rindfleisch, 2005), early family environment (Flouri, 2004), and peer influences (Roberts, Manolis, and Tanner, 2009). The developmental progression of materialistic tendencies in children has also been investigated (Chaplin and John, 2007).

Materialism is associated with many variables of interest to marketers, including a preference for status goods and unique products, the centrality of visual aesthetics when making a product choice, willingness to purchase counterfeit products, impulse buying, and other decision variables. It is also associated with a greater willingness to go into debt to purchase discretionary goods, lower levels of frugality, compulsive spending, and with perceived financial distress and conflict between spouses.

Bibliography

Belk, R.W. (1985) Materialism: trait aspects of living in the material world. *Journal of Consumer Research*, **12**, 265–280.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Chaplin, L.N. and John, D.R. (2007) Growing up in a material world: age differences in materialism in children and adolescents. *Journal of Consumer Research*, **34**, 480–493.

Christopher, A.N., Saliba, L., and Deadmarsh, E.J. (2009) Materialism and well-being: the mediating effect of locus of control. *Personality and Individual Differences*, **46**, 682–686.

Flouri, E. (2004) Exploring the relationship between mothers' and fathers' parenting practices and children's materialist values. *Journal of Economic Psychology*, **25**, 743–752.

Inglehart, R.F. (2008) Changing values among western publics from 1970 to 2006. *West European Politics*, **31**, 130–146.

Kasser, T. (2002) *The High Price of Materialism*, MIT Press, Cambridge.

Richins, M.L. (2004) The material values scale: a re-inquiry into its measurement properties and the development of a short form. *Journal of Consumer Research*, **31**, 209–219.

Richins, M.L. and Dawson, S. (1992) A consumer values orientation for materialism and its measurement: scale development and validation. *Journal of Consumer Research*, **19**, 303–316.

Roberts, J.A., Manolis, C., and Tanner, J.F. Jr. (2009) Interpersonal influence and adolescent materialism and compulsive buying. *Social Influence*, **3**, 114–131.

Shrum, L.J., Burroughs, J.E., and Rindfleisch, A. (2005) Television's cultivation of material values. *Journal of Consumer Research*, **32**, 473–479.

consumer memory processes

Kathryn R. Mercurio and Mark R. Forehand

Consumer memory involves the encoding, storage, and retrieval of information related to products and services. The strongest consumer brands are those that have high brand awareness, whereby consumers easily recognize the brand (aided awareness) and can also recall the brand in the absence of cues (unaided awareness). Following associative network models of memory (Anderson and Bower, 1973), consumer memory is generally described as a massive network of interconnected ideas and concepts. The probability that any particular piece of information will be retrieved is dependent on the strength of association between the information and other connected concepts that are activated in the environment. When numerous pieces of information are integrated into a given associative network, the probability of successful recall of any particular piece of information may drop. For example, memory performance often lessens in the presence of competing information from competitive advertising (Keller, 1991) or from contextual interference (Kumar and Krishnan, 2004).

Consumer memory performance is critically influenced by the depth with which new information is encoded and the context in which the encoding takes place. For example, brands that are encoded in conjunction with parent categories and are considered more prototypical members of a category are usually easier to recall than are less representative brands (Mao and Krishnan, 2006). Brand concept maps have shown that the core brand associations that define a brand's image are connected not only to brand but also to secondary associations outside the brand (John *et al.*, 2006). These secondary associations provide an initial indication of the concepts that are likely to trigger memories for the brand in the external environment. There are also a number of marketing-controlled factors at the time of encoding that influence memory performance including message source (Kirmani and Shiv, 1998), persuasion (Drolet and Aaker, 2002), order effects (Cunha, Janiszewski, and Laran, 2008), virtual product experience (Schlosser, 2006), sponsorship (Cornwell *et al.*, 2006), mode of encoding – visually, verbally, and so on (Vanhuele, Laurent, and Drèze, 2006), and repetition (Appleton-Knapp, Bjork, and Wickens, 2005).

The states a consumer is in during encoding and retrieval can also significantly influence the consumer's ability to retrieve learned information. Within an associative network, activated states can serve as central units or nodes in a semantic network and a reoccurrence of a state can, therefore, trigger improved recall of information associated with that state (Cowley, 2007). Two accepted models of state-based memory are state dependency and state congruency. State-dependent learning models suggest that memory performance improves when information learned under a particular activated state is retrieved while that state is again activated, regardless of whether the learned information is related to the state. In contrast, state congruent models propose that memory performance is driven by a match between an individual's state at retrieval and the a priori association of the learned content with the activated state (Fiedler, 2000). In advertising contexts that feature content ambiguously associated with a state, memory performance is only apparent when the same state is active at encoding and retrieval and the content is moderately related to the state (Mercurio and Forehand, 2009).

One way of improving memory performance in consumer contexts is to provide retrieval cues that activate critical information nodes that can in turn facilitate retrieval of closely associated information. Although these cues generally improve memory performance, they can also hinder long-term memory if the consumer becomes dependent on the cue for retrieval and does not reinforce the required retrieval pathways (Forehand and Keller, 1996).

Consumer memory is a dynamic and complex process and given the role of memory in almost all aspects of consumer behavior it is important to understand the marketing factors that facilitate or hinder memory performance. Memory performance depends on the ability to encode information into long-term memory, retain the information over time, and retrieve previously stored memory. As shown by the associative

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

network model of memory, factors present at encoding and retrieval can influence the strength of association between brands and new information and thereby significantly influence subsequent recall and recognition of consumer information.

Bibliography

Anderson, J.R. and Bower, G.H. (1973) *Human Associative Memory*, Winston, Washington, DC.

Appleton-Knapp, S.L., Bjork, R.A., and Wickens, T.D. (2005) Examining the spacing effect in advertising: encoding variability, retrieval processes, and their interaction. *Journal of Consumer Research*, **32**, 266–276.

Cornwell, B.T., Humphreys, M.S., Maguire, A.M. *et al*. (2006) Sponsorship-linked marketing: the role of articulation in memory. *Journal of Consumer Research*, **33**, 312–321.

Cowley, E. (2007) How enjoyable was it? Remembering an affective reaction to a previous consumption experience. *Journal of Consumer Research*, **34**, 495–505.

Cunha, M.V., Janiszewski, C., and Laran, J. (2008) Protection of prior learning in complex consumer learning environments. *Journal of Consumer Research*, **34**, 850–864.

Drolet, A. and Aaker, J. (2002) Off-target? Changing cognitive-based attitudes. *Journal of Consumer Psychology*, **12**, 59–68.

Fiedler, K. (2000) Towards and integrative account of affect and cognition phenomena using the BIAS computer algorithm, in *Feeling and Thinking: The Role of Affect in Social Cognition* (ed. J.P. Forgas), Cambridge University Press, New York.

Forehand, M.R. and Keller, K.L. (1996) Initial retrieval difficulty and subsequent recall in an advertising setting. *Journal of Consumer Psychology*, **5**, 299–323.

John, D.R., Loken, B., Kim, K., and Monga, A.B. (2006) Brand concept maps: a methodology for identifying brand association networks. *Journal of Marketing*, **43**, 549–563.

Keller, K.L. (1991) Memory and evaluation effects in competitive advertising environments. *Journal of Consumer Research*, **17**, 463–476.

Kirmani, A. and Shiv, B. (1998) Effects of source congruity on brand attitudes and beliefs: the moderating role of issue relevant elaboration. *Journal of Consumer Psychology*, **7**, 25–47.

Kumar, A. and Krishnan, S. (2004) Memory interference in advertising: a replication and extension. *Journal of Consumer Research*, **30**, 602–611.

Mao, H. and Krishnan, H.S. (2006) Effects of prototype and exemplar fit on brand extension evaluations: a two-process contingency model. *Journal of Consumer Research*, **33**, 41–49.

Mercurio, K.R. and Forehand, M.R. (2009) An interpretive frame model of state dependent learning: the moderating role of content – state association. *Journal of Consumer Research*, under review.

Schlosser, A.E. (2006) Learning through virtual product experience: the role of imagery on true and false memories. *Journal of Consumer Research*, **33**, 377–383.

Vanhuele, M., Laurent, G., and Drèze, X. (2006) Consumers' immediate memory for prices. *Journal of Consumer Research*, **33**, 163–172.

online consumption

David Mazursky and Gideon Vinitzky

In view of its fast growth, little doubt remains about the central role that online consumption plays in practically every aspect of our marketing world. Much of the research interest in recent years has focused on whether the key consumer decision-making rules that we acquired over decades of research can be easily applied in the online consumption context, or alternatively, whether it involves unique characteristics which challenge our accumulated knowledge and predictions about how consumers behave. The emerging literature indicates that the on-line shopping process possesses several unique characteristics that distinguish it from other forms of transactions.

The study of on-line shopping focuses mainly on the effects of technological aspects of the Internet on consumers' experience, motivation to engage in the process, the unique process dynamics, and postconsumption behavior. The literature developed so far mentions two technological features of the Internet as affecting the way people make their purchases and the outcome of consumption: interactivity and vividness (Steuer, 1992). Although the literature offers different definitions of the interactivity concept (Rafaeli, 1990; Steuer, 1992) and different dimensions based on various theoretical frameworks, it is agreed that most web sites offer dimensions that enable consumers to hold instant-response dialogues with the electronic interface and salesperson, simulating a regular conversation between human beings. The second characteristic of the technological environment is vividness, defined as the clarity of information received by consumers in the virtual world (Steuer, 1992).

Researchers suggest that interactivity and vividness promote the experience of telepresence, defined as consumers' feeling when located within the Internet environment (Steuer, 1992). Telepresence enables consumers to enjoy the actual process of navigating the site and is characterized by intrinsic motivation. Hoffman and Novak (1996) suggest that the interaction between the consumer and the purchase interface has an effect beyond that of the information in the shop, thus promoting the experience of flow, that is, the complete engagement with and immersion in an activity (Hoffman and Novak, 2009). The research states that the flow experience is more a characteristic of consumers in purchase tasks than of computer users on surfing tasks who have no purchase intentions and that the flow experience promotes exploratory, curiosity, and discovery behavior (control and perceived behavioral control, and learning).

What motivates consumers to engage in on-line shopping? From the financial aspect, buying on-line enables consumers to visit a large number of shops, increases the choice range of various products, increases competitiveness and gains from reduced prices (Bakos, 1997; Lynch and Ariely, 2000). From the cognitive aspect, consumers expect to benefit from the reduced cost of cognitive search (Lurie, 2004; Häubl and Trifts, 2000; Alba *et al.*, 1997), affecting also their sensitivity to price and quality (Lynch and Ariely, 2000). From the hedonic perspective, researchers suggest that on-line transactions provide consumers with types of pleasure that are similar to traditional transactions (Menon and Kahn, 2002) as well as other sources of enjoyment, for example, convenience and entertainment, that are uniquely associated with on-line purchasing (Childers *et al.*, 2001).

One of the most significant barriers to consumer entry when purchasing on-line is the concern of trust. Trust is a psychological state expressing reliance on another person or organization to do that what is required. Researchers state that trust is a multidimensional psychological concept that includes cognition, affect, and behavior (Johnson and Grayson, 2005). More concretely, factors affecting trust include site features (Shankar and Sultan, 2002), familiarity with the company and its reputation (Yoon, 2002), information security features (Belanger, Hiller, and Smith, 2002), ease of use, and expertise (Fogg *et al.*, 2001). At the same time, trust-motivating factors are distinctively affected by various product categories and interpersonal differences (Bart *et al.*, 2005).

Research on the dynamics of the on-line shopping process suggests that consumers demonstrate various search patterns stemming from the features of the interface (Mazursky and Vinitzky, 2005). Three-dimensional purchase

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

interfaces were found to promote a structured search process, which is more continuous and more compatible with physical shopping, while two-dimensional purchase interfaces were found to promote less structured search, which is quicker yet different from the search pattern characteristic of buying in a traditional shop.

On the basis of the preference construction approach (Bettman, Luce, and Payne, 1998), researchers suggest that consumer preferences change during the purchase process in line with various aspects of the purchase environment. Häubl and Trifts (2000) propose that the default values that appear in search agents affect the features examined in the purchase process. Mandel and Johnson (2002) found that the background color of a web site can serve as priming the criteria involved in product selection and that it affects the weight given to these criteria in the process of product choice.

The postconsumption factors consist of satisfaction, loyalty, and word-of-mouth communication. Studies offer various models for examining consumer satisfaction upon purchasing on-line. Szymanski and Hise (2000) suggest that the surfer's convenience, the products offered, site design, and financial safety issues, have an effect on satisfaction. Cheung and Lee (2005) focused on the quality of the various components containing the information displayed in the shop and how the system functions. Turban *et al.* (2008) expanded this model and added the component of quality of service provided on site. Surfer satisfaction, in turn, was found to impinge on a variety of shopping related aspects, including word-of-mouth (Parish, Holloway, and Wang, 2005; Chevalier and Mayzlin, 2006), trust (Bart *et al.*, 2005), and loyalty (Bolton, Kannan, and Bramlett, 2000; Floh and Treiblmaier, 2006).

The ease of access to a variety of alternative shops offering the same services (Lynch and Ariely, 2000) contributes positively to consumer loyalty. Keaveney and Parthasarathy (2001) suggest that previous consumer behavior patterns on the site (e.g., use of services), attitudes, involvement, and demographic variables, distinguish between consumers' loyalty status. Other factors found to promote loyalty are the quality of services on the site (Chen and Hitt, 2002), consumer commitment (Pavlou and Gefen, 2004), and trust (Park and Kim, 2006).

However, experiences in on-line shopping may not always be positive. Presumably because of the relative ease to express negative feedback (in comparison with other forms of transaction), such activity is highly prevalent and has intrigued researchers' interest. The nonpositive pattern of postshopping behavior may involve revenge, negative rumor, and negative word of mouth. For example, writing negative on-line reviews may even lead to harming of sales (e.g., Chevalier and Mayzlin, 2006)

Finally, despite the noticeable progress in the study of on-line shopping behavior, its unique infrastructure represents an important research area both for studying on-line shopping, as well as for serving as a proxy for learning the way consumers form decisions in shopping behavior, in general. This infrastructure involves presenting information, prompting a decision, and echoing postconsumption reactions, all in vivo, and in a relatively condensed time framework, enabling tracking and analyzing the process. This is quite a unique setting compared with other forms of transactions such as in physical stores, where tracking this process generally involves tools that are largely artificial and inaccurate. In addition, the growing direct linking possibilities from the shopping activity to other network systems such as social networks, available in on-line shopping formats, further increase their research potential. The outlook is indeed promising, although possible barriers such as intercultural and language issues need to be carefully considered in planning future research.

See also *customer satisfaction; e-commerce and Internet marketing; marketing functions on the Internet; social networks*

Bibliography

Alba, J.W., Lynch, J., Weitz, B. *et al.* (1997) Interactive home shopping: consumer, retailer, and manufacturer incentives to participate in electronic marketplaces. *Journal of Marketing*, **61**, 38–53.

Bakos, J.Y. (1997) Reducing buyer search cost: implication for electronic marketplaces. *Management Science*, **43** (12), 1676–1693.

Bart, Y., Shankar, V., Sultan, F. and Urban, G.L. (2005) Are the drivers and role of online trust the same for all web sites and consumers? A large-scale exploratory empirical study. *Journal of Marketing Research*, **69** (4), 133–152.

Belanger, F., Hiller, J.S. and Smith, W.J. (2002) Trustworthiness in electronic commerce: the role of privacy, security, and site attributes. *Journal of Strategic Information Systems*, **11** (3,4), 245–270.

Bettman, J.R., Luce, M.F. and Payne, J.W. (1998) Constructive consumer choice processes. *Journal of Consumer Research*, **25** (3), 187–217.

Bolton, R.N., Kannan, P. and Bramlett, M.D. (2000) Implications of loyalty program membership and service experiences for customer retention and value. *Journal of the Academy of Marketing Science*, **28** (1), 95–108.

Chen, P. and Hitt, L. (2002) Measuring switching costs and the determinants of customer retention in Internet-enabled businesses: a study of the online brokerage industry. *Information Systems Research*, **13** (3), 255–274.

Cheung, C.M.K. and Lee, M.K.O. (2005) The Asymmetric Impact of Website Attribute Performance on User Satisfaction: An Empirical Study. Paper presented at the Hawaii International Conference on System Sciences, Big Island, Hawaii.

Chevalier, J.A. and Mayzlin, D. (2006) The effect of word of mouth on sales: online book reviews. *Journal of Marketing Research*, **43** (3), 9.

Childers, T.L., Carr, C.L., Peck, J. and Carson, S. (2001) Hedonic and utilitarian motivations for online retail shopping behavior. *Journal of Retailing*, **77** (4), 511.

Johnson, D. and Grayson, K. (2005) Cognitive and affective trust in service. Relationships. *Journal of Business Research*, **58** (4), 500–507.

Floh, A. and Treiblmaier, H. (2006) What keeps the e-banking customer loyal? A multigroup analysis of the moderating role of consumer characteristics e-loyalty in the financial service industry. *Journal of Electronic Commerce Research*, **7** (2), 97–110.

Fogg, B.J., Marshall, J., Laraki, O. *et al*. (2001) What makes web sites credible? A report on a large quantitative study. *ACM SIGCHI*, **3** (1), 61–67.

Häubl, G. and Trifts, V. (2000) Consumer decision making in online shopping environments: the effects of interactive decision aids. *Marketing Science*, **19** (1), 4–21.

Hoffman, D.L. and Novak, T.P. (1996) Marketing in hypermedia computer- mediated environment: conceptual foundation. *Journal of Marketing*, **60**, 50–68.

Hoffman, D.L. and Novak, T.P. (2009) Flow online: lessons learned and future prospects. *Journal of Interactive Marketing*, **23** (1), 23–34.

Keaveney, S. and Parthasarathy, M. (2001) Customer switching behavior in online services: an exploratory study of the role of selected attitudinal, behavioral, and demographic factors. *Journal of the Academy of Marketing Science*, **29** (4), 374–390.

Lurie, N.H. (2004) Decision making in information-rich environments: the role of information structure. *Journal of Consumer Research*, **30** (4), 473–486.

Lynch, J. and Ariely, D. (2000) Wine online: search costs affect competition on price, quality and distribution. *Marketing Science*, **19**, 83–103.

Mandel, N. and Johnson, E.J. (2002). When web pages influence choice: effects of visual primes on experts and novices. *Journal of Consumer Research*, **29** (2), 235–245.

Mazursky, D. and Vinitzky, G. (2005) Modifying consumer search processes in enhanced on-line interfaces. *Journal of Business Research*, **58** (10), 1299.

Menon, S. and Kahn, B. (2002) Cross-category effects of induced arousal and pleasure on the Internet Shopping Experience. *Journal of Retailing*, **78** (1), 31–40.

Parish, J.T., Holloway, B.B. and Wang, S. (2005) The role of cumulative online purchasing experience in service recovery management. *Journal of Interactive Marketing*, **19** (3), 54–66.

Park, C. and Kim, Y. (2006) The effect of information satisfaction and relational benefit on consumers' online shopping site commitments. *Journal of Electronic Commerce in Organizations*, **4** (1), 70–90.

Pavlou, P. and Gefen, D. (2004) Building effective online marketplaces with institution-based trust. *Information Systems Research*, **15** (1), 37–59.

Rafaeli, S. (1990). Interacting with media: para-social interaction and real interaction, in *Mediation, Information and Communication: Information and Behavior* (ed. B.D. Ruben and L.A. Lievrouw), pp. 125–181, Vol. 3.

Shankar, G.L.U. and Sultan, F. (2002) Online trust: a stakeholder perspective, concepts, implications and future directions. *Journal of Strategic Information Systems*, **11** (3-4), 325–344.

Steuer, J. (1992) Defining virtual reality dimensions determining telepresence. *Journal of Communication*, **42** (4), 73–93.

Szymanski, D.M. and Hise, R. (2000) E-satisfaction: an initial examination. *Journal of Retailing*, **76** (3), 309–322.

Turban, E., Lee, J.K., KIng, D. *et al*. (2008) *Electronic Commerce*, 1st edn, Prentice Hall, New Jersey.

Yoon, S.-J. (2002) The antecedents and consequences of trust in online purchase decisions. *Journal of Interactive Marketing*, **16** (2), 47–63.

consumer acculturation

Lisa Peñaloza

Consumer acculturation is defined as the general process of adaptation by consumers in the marketplace. While initially theorized for immigrants seeking economic opportunity and quality of life in another nation or fleeing situations of economic conflict or war at home (Peñaloza, 1994), distinct consumption patterns and market relations are forged by minority groups as well, as they adapt to mainstream and other groups in heterogeneous societies (Kjeldgaard and Askegaard, 2006). Such interaction and adaptation is increasingly widespread globally, as families and communities extend beyond national borders. The growing importance of the topic is evident from the 80 odd participants of over 20 nationalities at the May 2009 conference on Immigration, Consumption and Markets organized by Nil Toulouse and Soren Askegaard in Lille, France.

The scope of this body of work is broad, in examining social and market learning and interaction at individual, family, community, societal, and transnational levels (Lindridge, Hogg, and Shah, 2004; Askegaard, Arnould, and Kjeldgaard, 2005; Parreñas, 2005; Nolin, 2006; Peñaloza, 2007). Patterns of adaptation include a varied, dynamic mix of assimilating new consumption tastes and skills, maintaining old ones, resisting new and old forms, swapping or alternating between them (Oswald, 1999; Visconti, 2005), and forming distinctively hybrid consumption patterns based on localized interpretations of global media and market artifacts (Kjeldgaard and Askegaard, 2006) within global labor (Peñaloza, 1995) and financial imperatives (Üçok, 2007).

Consumer research challenges include distinguishing consumption practices across groups and calibrating changes in cultural skills, values, and knowledge, as inflected by social class, race, ethnicity, gender, religion, nationality, and geographic differences, and as played out in increasingly diverse intra- and transnational social and market formations. As cultural artifacts and behaviors are detached from the people creating them and appropriated by other consumers and corporations globally (Grier, Brumbaugh, and Thornton, 2006), concerns of cultural hegemony and homogenization among minority groups coincide with concerns with loss of status and power by the mainstream group, as both seek to maintain cultural identity and integrity. These concerns gain currency with the decreasing size of mainstream cultural groups and as minority consumer subcultures are incorporated by marketers within nations and in market diaspora across the globe.

Work on marketer acculturation examines how marketers learn to serve consumers of various cultures and address them in ways sensitive to relations between groups (Peñaloza and Gilly, 1999). Carried further, acculturation research can provide insight into the identity and vitality of nations by examining the nature of group relations as diverse people coexist and adapt their respective ways of life in consumption, in markets, and in communities. The issue is the role of business in separating, uniting, legitimizing, and excluding groups of people, as tailoring products and services to the mainstream advances, its dominance, even as targeting groups outside a cultural mainstream enables their members to reproduce vital cultural patterns and meanings and thrive. Public policies restricting the language and other cultural expressions of those who are "different" or "other" and excluding them in markets and other social services are the means by which one group dominates the others (Üstüner and Holt, 2007). Such intolerance increases in adverse economic circumstances and may be accompanied by violence, and yet such intolerance can trigger politically oriented consumption and market incorporation that ultimately advances the identity and community of those otherwise excluded (Peñaloza, 2007).

Further research opportunities lie in documenting how market development relies on and fosters forms of difference. Ultimately, the market both integrates and separates (Keating and McLoughlin, 2005). In their integrative function, markets assemble various people under a particular rubric and provide them with the cultural legitimization of a market. Other promising areas of work are investigating the nature of market participation played out within cities and nations, and exploring how immigrants and other minorities attain forms of

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

consumer sovereignty that potentially inform by and impact political participation. It remains to be seen how capitalism and democracy will continue to cross-fertilize each other, with immigrants and minorities serving as test cases.

Bibliography

Askegaard, S., Arnould, E.J., and Kjeldgaard, D. (2005) Postassimilationist ethnic consumer research: qualifications and extensions. *Journal of Consumer Research*, **32**, 160–170.

Grier, S., Brumbaugh, A.M., and Corlis, T. (2006) Crossover dreams: consumer responses to ethnic-oriented products. *Journal of Marketing*, **70**, 35–51.

Keating, A. and McLoughlin, D. (2005) Understanding the emergence of markets: a social constructionist perspective on gay economy. *Consumption, Markets, Culture*, **8**, 131–152.

Kjeldgaard, D. and Askegaard, S. (2006) The glocalization of youth culture: the global youth segment as structures of common difference. *Journal of Consumer Research*, **33**, 231–247.

Lindridge, A., Hogg, M., and Shah, M. (2004) Imagined multiple worlds: how South Asian Women in Britain use family and friends to navigate the border crossings between household and societal contexts. *Consumption, Markets, Culture*, **7**, 211–238.

Nolin, C. (2006) *Transnational Ruptures. Gender and Forced Migration*, Ashgate, London.

Oswald, L. (1999) Culture swapping: consumption and ethnogenesis of middle class hatian immigrans. *Journal of Consumer Research*, **25**, 303–318.

Parreñas, R.S. (2005) *Children of Global Migration: Transnational Families and Gender Woes*, Stanford University Press, Stanford, CA.

Peñaloza, L. (1994) Atravasando fronteras/border crossings. *Journal of Consumer Research*, **21**, 32–54.

Peñaloza, L. (1995) Immigrant consumers: marketing and public policy considerations in the global economy. *Journal of Public Policy and Marketing*, **14**, 83–94.

Peñaloza, L. (2007) Mainstreet U.S.A revisited: market targeting, Latino/a consumer culture, and community. *International Journal of Sociology and Social Policy*, **27**, 234–249.

Peñaloza, L. and Mary G. (1999) arketers' acculturation: the changer and the changed. *Journal of Marketing*, **63**, 84–104.

Üçok, M. (2007) Consumption practices in transnational social spaces: a study of Turkish transmigrants in Denmark. Doctoral dissertation, University of Southern Denmark, Odense.

Üçokner, T. and Holt, D.B. (2007) Dominated consumer acculturation: the social construction of poor migrant women's consumer identity projects in a Turkish squatter. *Journal of Consumer Research*, **34**, 41–56.

Visconti, L. (2005) Attraversate le frontiere/border crossed: processes of cultural alternation in the marketplace. Unpublished Ph.D. dissertation, Management Department, Bocconi University, Milan, Italy.

subcultures

Robert V. Kozinets

Contemporary consumer culture is filled with a rich gallery of subcultural archetypes, including rappers, gangstas, punks, surfers, jocks, fanboys, goth girls, metalheads, Trekkers, and hackers. These subcultures provide meanings and practices that significantly structure consumers' identities, relationships, and behaviors in many ways. They offer up rich, meaningful, group identities that differentiate from, oppose, or even seek to subvert normal mainstream ways of speaking, thinking, and behaving. With their resistant, activist, symbolic characteristics, subcultures are closely related to countercultures and often serve as the source of countercultural images and identities. These images and identities are often absorbed by the advertising, entertainment, and marketing industries, and eventually become popularized and commonplace mainstream activities. Subcultures thus fuel the never-ending engine of differentiation that underlies contemporary consumer culture (Heath and Potter, 2004).

In consumer research or consumer cultural studies, subcultures have been related to sports and avocations, fantasy and escape, mass media fans, brand consumption, as well as gender and class (Belk and Costa, 1998; Celsi, Rose, and Leigh, 1993; Kozinets, 2001; Schouten and McAlexander, 1995). Different aspects of consumption-related subcultures have been emphasized by consumer researchers, including their mode of acculturation (Celsi, Rose, and Leigh, 1993), their self-selection and hierarchical, ethos-driven structure (Schouten and McAlexander, 1995), as well as their relationship to mass-mediated meanings and images, resistance, and activism, and online communities (Kozinets, 2001, 2002; Kozinets and Handelman, 2004).

John Schouten and Jim McAlexander (1995, p. 43) coined the useful term *subculture of consumption* to refer to and underscore the importance of these consumption-oriented groupings, defining the term as "a distinctive subgroup of society that self-selects on the basis of a shared commitment to a particular product class, brand, or consumption activity." They also summarized and suggested some key characteristics of these groups: their shared ethos, mythos, values, and beliefs; their social structure based on hierarchies of commitment and authenticity; their shared language and rituals; and the religious and transformational aspects of the subcultures. However, linking consumer culture to the subcultures literature (e.g., Hebdige, 1979) has drawn some criticism. Some argue that "the social groups investigated in the name of 'subcultures' are subordinate, subaltern or subterranean" or are "deviant," "debased," illegitimate or of lower socioeconomic status" (Thornton, 1997, p. 4) or that the term is ambiguous and negatively charged (Hannerz, 1992). A further critique holds that the term has been overextended to refer to a variety of different leisure activities and hobbies, rather than to actual ways of life or social forms. In response to these and related critiques in sociology and anthropology, a number of other conceptions, including "cultures of consumption" (Kozinets, 2001), "microcultures" (Hannerz, 1992), "brand communities" (Muniz and O'Guinn, 2001, *see* BRAND COMMUNITY), and "consumer tribes" (Cova, Kozinets, and Shankar, 2007) have been proposed and used.

Yet despite the ever-present ambiguities of the term, the concept of subcultures is still very useful to consumer researchers and others who are interested in understanding consumer behavior in its actual, lived sense. The term deals with an ambiguous, dynamic, and countercultural phenomenon with a long history, strong momentum, and deep ties to consumer culture. The underground connotations of the term *subculture*, its established links to foundational cultural studies literature, and its inherent flexibility will continue to inspire new generations of researchers for many years to come. As long as lifestyle, differentiation, self-transformation, and small group identities continue to play an important part of consumers' lives in contemporary consumer culture, the concept of subcultures will help us to better understand the complex social world around us.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Bibliography

Belk, R.W. and Costa, J.A. (1998) The mountain man myth: a contemporary consuming fantasy. *Journal of Consumer Research*, **25**, 218–240.

Celsi, R.L., Rose, R.L., and Leigh, T.W. (1993) An exploration of high-risk consumption through skydiving. *Journal of Consumer Research*, **20**, 1–23.

Cova, B., Kozinets, R.V., and Shankar, A. (eds) (2007) *Consumer Tribes*. Butterworth-Heinemann, Oxford and Burlington.

Hannerz, U. (1992) *Cultural Complexity*, Columbia University, New York.

Heath, J. and Potter, A. (2004) *Nation of Rebels: Why Counterculture became Consumer Culture*, Harper Business, New York.

Hebdige, D. (1979) *Subculture*, Methuen, New York.

Kozinets, R.V. (2001) Utopian enterprise: articulating the meanings of Star Trek's culture of consumption. *Journal of Consumer Research*, **28**, 67–88.

Kozinets, R.V. (2002) The field behind the screen: using netnography for marketing research in online communities. *Journal of Marketing Research*, **39**, 61–72.

Kozinets, R.V. and Handelman, J.M. (2004) Adversaries of consumption: consumer movements, activism, and ideology. *Journal of Consumer Research*, **31**, 691–704.

Muñiz, A.M. Jr. and O'Guinn. T.C. (2001) Brand community. *Journal of Consumer Research*, **27**, 412–432.

Schouten, J.W. and McAlexander, J.H. (1995) Subcultures of consumption: an ethnography of the new bikers. *Journal of Consumer Research*, **22**, 43–61.

Thornton, S. (1997) General introduction, in *The Subcultures Reader* (eds K.Gelder and S., Thornton), Routledge, New York, pp. 1–7.

social class

David K. Crockett

What is Social Class?

Although no single precise definition exists, most scholars acknowledge that social class is composed of a variety of factors such as wealth (especially inherited wealth), income, occupational status, educational attainment, and residence. Arnould, Price, and Zinkhan (2002, p. 181) provide a useful working definition: *social classes are groupings across society, broadly recognized by members of that society, involving inequalities, or certainly, differences in such areas as power, authority, wealth, income, prestige, working conditions, lifestyles, and culture*. Social class is by nature hierarchical, that is, groups and individuals receive unequal amounts of power, prestige, and esteem (i.e., status). Social class is pervasive, if often invisible. The mere mention of it may invite feelings of superiority among members of the elite, or feelings of uneasiness among those at the bottom of the status hierarchy.

One important component of social class is income. In fact many people think income is synonymous with class. However, even though income largely determines lifestyle at any given point in time it is a poor indicator of social class. In much of the developing world, accurate income data is difficult to obtain, and in developed nations income can fluctuate wildly over the life course (Leonhardt, 2005). Hence, instead of income many organizations and institutions use *purchasing power parities* to make cross-national comparisons where costs of living and currency values can be quite different.[1]

How does Social Class Impact Consumer Behavior?

Scholars have historically sought one-to-one relationships between income (usually indexed with education and occupation) and consumption, where high status consumers purchase brand X while lower status consumers purchase comparably more brand Y. In contemporary consumer culture however, the lowest and highest status consumers often choose the same products (Schor, 1998). Consequently, the old economic approaches are giving way to cultural approaches that generate insights about social class and consumption by focusing on *how* people consume rather than *what* they consume (Holt, 1998). People consume their way up the social class hierarchy by utilizing three types of status-generating resources: economic capital, social capital, and cultural capital (Bourdieu, 1984). Cultural approaches emphasize social and cultural capital (while economic approaches emphasize only economic capital). Social capital refers to membership in formal and informal reference groups. Cultural capital refers to skills, knowledge, tastes, and preferences derived indirectly from family background, class heritage, and common experience. Higher status comes from consuming in ways that signal a person's store of social and cultural capital relative to others.

What are the Social Class Segments?

The particular distribution of individuals into social class groupings varies widely from country to country. Most countries feature at least three segments (upper, middle, and working class), with the most fully formed class structures in the developed world. The upper class can include upper-upper and lower-upper subsegments. The former is usually characterized by inherited wealth, while the latter is often characterized by "nouveau riches" (or new money). Many, if not most consumption trends are indexed to the tastes of these two subsegments (Schor, 1998). The upper middle class, despite lower stores of economic capital than the upper class, is typically more similar to the upper classes than to other members of the middle class in aspirational lifestyles, culture, and values. The remainder of the middle class is the primary target of most marketing efforts aimed directly at consumers. The majority of the world's people, however, belong to the working class or the chronically poor. Working-class consumers are difficult to characterize globally, other than by their very limited economic resources. In developed nations, they are thought to basically enjoy the core features of their current lifestyles and have little incentive to fundamentally alter them (other than to have more resources). Finally, the chronically poor or underclass largely lack

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

access to most basic consumer choices. World Bank researchers estimate that one-quarter of the developing world fits in this group, living on less than $1.25 per day in 2005 prices (Chen and Ravaillion, 2008).

Endnotes

[1] The World Bank estimates purchasing power parities between countries (where data is available) and publishes the data at its website: http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/ICPEXT/0,,pagePK:62002243~theSitePK:270065,00.html

See also *consumer acculturation; consumer behavior across literacy and resource barriers; consumer materialism; consumer well-being; global consumerism and consumption; social class; social influence; social networks; society, culture, and global consumer culture; subcultures*

Bibliography

Arnould, E.J., Price, L.L. and Zinkhan, G. (2002) *Consumers*, McGraw Hill, New York.

Bourdieu, P. (1984) *Distinction: A Social Critique of the Judgment of Taste*, Harvard University Press, Cambridge, MA.

Chen, S. and Ravaillion, M. (2008) The Developing World Is Poorer Than We Thought, But No Less Successful in the Fight against Poverty, *Policy Research Paper #4703*, World Bank Development Research Group, http: //econ. worldbank. org/external/default/main?pagePK=64165259&theSitePK=469372&piPK=64165421&menuPK=64166093&entityID=000158349_20080826113239] (accessed 31 May 2009).

Holt, D.B. (1998) Does cultural capital structure american consumption? *Journal of Consumer Research*, **25** (1), 1–25.

Leonhardt, D. (2005) A closer look at income immobility. (May 14) *New York Times* (online), http://www.nytimes.com/2005/05/14/national/class/15MOBILITY-WEB.html?_r=1] (accessed 31 May 2009).

Schor, J.B. (1998) *The Overspent American: Upscaling, Downshifting, and the New Consumer*, Basic Books, New York.

impulsive and compulsive buying

Ronald J. Faber

Impulsive and compulsive buying are terms that are frequently confused for each other, but represent behaviors that differ greatly in regard to their frequency, cause, outcome, and severity. Impulsive buying is a more common and ordinary behavior. Virtually everyone makes a purchase on impulse (without much deliberation) from time to time. Some do this more frequently than others. Nonetheless, almost everyone finds that sometimes they are able to resist the desire to buy on impulse and at other times they give in to this impulse. It has been suggested that impulse buying occurs when desire for a product or brand outweighs one's willpower to resist (Hoch and Loewenstein, 1991).

On the other hand, compulsive buying is a psychological disorder where one experiences an urge to buy that cannot be controlled. Failing to act on this urge creates increasing tension for the individual which only dissipates with buying. Frequently, this urge is triggered by negative events or feelings. Ultimately, this behavior leads to extreme negative consequences (financial and/or personal) for the individual (Faber and Christenson, 1996; O'Guinn and Faber, 1989). Many compulsive buyers never use the items they purchase.

Impulse Buying

Historical development of impulse buying. Early research on impulse buying in the 1950s and 1960s viewed it as occurring anytime a consumer made any unplanned purchase. As a result, research sought to classify certain types of products as impulse items and looked at what store characteristics encouraged impulse buying. This approach changed in the 1970s as researchers began to realize that all products could be purchased impulsively. As a result, research on impulse buying began to focus on characteristic of individuals, rather than of products or stores, and to define the term as a sudden and powerful urge in the consumer to buy immediately (Rook and Hoch, 1985; Rook, 1987).

Correlates and causes of impulse buying. Factors that influence both the desire for a product and willpower to resist this desire are related to situations when impulse buying occurs (Hoch and Loewenstein, 1991). Proximity to a purchase can increase wanting for an item. Thus one technique to reduce impulse buying is to walk away from an item and only buy it if you still want it as much several minutes later. Another factor that correlates with impulse buying is mood state. People indicate they are much more likely to buy on impulse when in a positive mood (Rook and Gardner, 1993).

Willpower has also been shown to effect impulse buying. Vohs and Faber (2007) showed that people were more likely to buy on impulse when their self-regulatory resources (willpower) were diminished. Engaging in self-regulation in some other domain increases the likelihood people will buy on impulse. Thus, not shopping after a long, difficult day and keeping shopping trips short can help people to cut down on impulse buying.

Compulsive Buying

Historical development of compulsive buying. Compulsive buying was initially discussed in the psychiatric literature under the term *onomania* almost 100 years ago (Kraeplin, 1915). However, there was little discussion of it from the mid-1920s until the late 1980s when researchers in the United States, Canada, and Germany all began to report on this problem (d'Astous, Maltais, and Roberge, 1990; O'Guinn and Faber, 1989; Scherhorn, Reisch, and Raab,1990). The Compulsive Buying Scale (Faber and O'Guinn, 1992) was developed to distinguish compulsive buyers from other consumers. Using this measure with a national probability sample, Koran *et al.* (2006) estimated that 5.8% of the US population may be compulsive buyers.

Correlates and causes of compulsive buying. Many compulsive buyers have personal or family histories of other problem behaviors such as alcoholism and drug abuse, eating disorders, and impulse control disorders. The most commonly reported correlates of compulsive buying are psychological ones. Numerous studies have found that compulsive buyers

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

are perfectionists, and have low self-esteem and high levels of depression and anxiety. Compulsive buying seems to provide temporary relief from these negative feelings, but ultimately makes compulsive buyers feel even worse about themselves. Thus, compulsive buying appears to be more about obtaining short-term relief from negative feelings than about a desire for specific goods. Several studies have found that compulsive buying is unrelated to income.

See also *consumer desire; consumer materialism; consumer well-being; emotion; possessions and self; self-regulation*

Bibliography

d'Astous, A., Maltais, J. and Roberge, C. (1990) Compulsive buying tendencies of adolescent consumers, in *Advances in Consumer Research*, vol. 17 (eds M.E. Goldberg, G. Gorn and R.W. Pollay), Association for Consumer Research, Provo, pp. 306–312.

Faber, R.J. and Christenson, G.A. (1996) In the mood to buy: differences in the mood states experienced by compulsive buyers and other consumers. *Psychology and Marketing*, **13**, 803–820.

Faber, R.J. and O'Guinn, T.C. (1992) A clinical screener for compulsive buying. *Journal of Consumer Research*, **19**, 459–469.

Hoch, S.J. and Loewenstein, G.F. (1991) Time inconsistent preferences and consumer self-control. *Journal of Consumer Research*, **18**, 492–507.

Koran, L., Faber, R.J., Aboujaoude, E. *et al*. (2006) Estimated prevalence of compulsive buying in the United States. *American Journal of Psychiatry*, **163** (10), 1806–1812.

Kraeplin, E. (1915) *Psychiatrie*, 8th edn, Verlag Von Johann Ambrosius Barth, Leipzig.

O'Guinn, T.C. and Faber, R.J. (1989) Compulsive buying: a phenomenological exploration. *Journal of Consumer Research*, **16**, 147–157.

Rook, D.W. (1987) The buying impulse. *Journal of Consumer Research*, **14**, 189–199.

Rook, D.W. and Gardner, M.P. (1993) In the mood: impulse buying's affective antecedents. *Research in Consumer Behavior*, **6**, 1–28.

Rook, D.W. and Hoch, S.J. (1985) Consuming impulses. *Advances in Consumer Research*, **12**, 23–27.

Scherhorn, G., Reisch, L.A. and Raab, G. (1990) Addictive buying in West Germany: an empirical study. *Journal of Consumer Policy*, **13**, 355–387.

Vohs, K. and Faber, R.J. (2007) Impulse buying: a result of self-regulatory resource depletion. *Journal of Consumer Research*, **33** (4), 537–548.

customer satisfaction

Richard L. Oliver

Introduction

This article describes the general field of consumer (customer) satisfaction behavior as it is currently conceptualized. Though the field has been researched from many perspectives, the most common appearing in the areas of satisfaction attribute (feature) *surveys* and customer satisfaction *strategy* (*see* CUSTOMER SATISFACTION/DISSATISFACTION), more psychologically based approaches are now practiced.

Here, the underlying mechanisms of how consumers construct, consciously or subconsciously, their satisfaction conclusions are explored so that greater satisfaction fostering and dissatisfaction avoidance can be ensured. Readers interested in greater detail and elaboration, including discussion of topics not covered here, should consult the author's original work (Oliver, 2010) and an earlier short treatise on satisfaction research (Oliver, 2006).

Customer Satisfaction Defined

Recent interpretations in the consumer domain now couch satisfaction as a fulfillment response. Fulfillment implies that a consumption goal is known, as in basic motives of hunger, thirst, and safety. However, observers of human behavior understand that these and other goals can be and frequently are modified and updated in various ways. Thus, consumer researchers have moved away from the traditional meaning of satisfaction and now pursue this concept as the consumer experiences and describes it.

In Oliver (2010, p. 8), the following definition has been proposed as being consistent with the conceptual and empirical evidence:

> Satisfaction is the consumer's fulfillment response. It is a judgment that a product or service feature, or the product or service itself, provided (or is providing) a pleasurable level of consumption-related fulfillment, including levels of under- or over-fulfillment.

Here, pleasurable implies that fulfillment gives pleasure or reduces pain. Thus, individuals can be satisfied so as to return to normalcy, as in the removal of an aversive state. Moreover, fulfillment is not necessarily limited to the case of met needs. Overfulfillment can be satisfying if it provides additional unexpected pleasure; and underfulfillment can be satisfying if it gives greater pleasure than one anticipates in a given situation. Note that if the word "displeasure" is substituted for pleasure in the satisfaction definition, dissatisfaction results. Thus, the displeasure of underfulfillment typically is dissatisfying and overfulfillment may be dissatisfying if it is unpleasant – "too much of a good thing."

A Satisfaction Model Useful for Current Thinking

At this point, it would be helpful to envision the framework upon which this discussion is based. If one can construct the antecedents of the satisfaction response, that is, how a consumer mentally constructs satisfaction, many conceptual nuances will be revealed. Perhaps the most useful model of this process is the expectancy disconfirmation framework that proceeds as follows. The model assumes that consumers have prior expectations of the product or service performance.

Expectations, whether measured before or after consumption (predicted or recalled retrospective expectations), and performance are compared to form an "objective" (or gap) disconfirmation level; objective disconfirmation provides the basis for a subjective interpretation of this expectation-performance difference, and subjective disconfirmation is one direct cause of satisfaction. Additionally, there may exist a direct effect of performance on satisfaction not channeled through disconfirmation. This represents the "expectancy disconfirmation with performance model."

The performance of features (attributes). In a familiar research scenario, consumers are asked to retrospectively rate the product or service on the degree to which each feature was delivered. Concurrently, the consumer may be asked to rate the product on an overall basis or on satisfaction.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Despite the ubiquity of this method, problems are inherent in its implementation. One is that the list of features cannot be exhaustive for all consumers. A second problem arises from the disparate goals (needs) of consumers. Another problem is that of the relevance of features at different stages of decision-making.

Satisfaction drivers versus choice criteria. In pursuing the reasons behind the consumer's satisfaction response, it should be borne in mind that the researcher's goal is to determine the correct feature list of *satisfaction drivers*, as opposed to product or service *choice criteria*. A common mistake is that of assuming the features consumers use in selecting a product from a list of alternatives are identical to the set of features that play into satisfaction and dissatisfaction judgments.

Thus, for this reason, customer satisfaction researchers (*see* CUSTOMER-SATISFACTION RESEARCH) are advised to determine satisfiers and dissatisfiers independently of choice determinants. Additionally, this also illustrates why a satisfaction measure is preferred to one of, say, attitude (*see* ATTITUDES) or quality, if consumer satisfaction is the goal of the firm. Both attitude and quality judgments are used in choice and thus may give a distorted picture of the features most strongly related to satisfaction.

Expectations and their role in satisfaction. Generally, an expectation is an anticipation of future consequences based on prior experience and other many and varied sources of information. Expectations can also be described as a *comparative referent* for performance. The reason is that performance alone is an unreferenced concept. Meaning is attached only when performance can be compared to some standard. In fact, any number of referents can be used in later satisfaction assessments, but they become channeled into expectations when the product or service is purchased. The concept of needs, discussed previously, is one of the many referents available to consumers. These same consumers, however, will pursue only those products that they *expect* to fulfill their needs.

Frequently, consumers express different variations of what they would prefer a product to deliver. At one level, they may have an ideal perception of a product offering, something they wish they could receive in a perfect world. Others, or the same consumers at a different time, expect only what they believe the firm's product can or will deliver. Researchers have referred to these two different perceptions as *ideal* and *predicted* expectations or, alternatively, as *should* and *will* or *desired* and *likely* outcomes.

A number of research investigations have established that consumers do indeed recognize and use multiple levels of expectations or standards. Among the most common of these are studies investigating the influence of normative (should) and predicted (will) expectations, or alternatively, ideal and expected referents. The results of all studies tend to be similar. When the ideal or should level of expectations was the referent, satisfaction was lower than when actual expected or predicted expectations were used. Moreover, these studies generally conclude that consumers do entertain multiple standards and that inclusion of more than just the predicted level may improve a model's ability to understand satisfaction.

Disconfirmation and its role in satisfaction. When consumers compare performance to their expectations, the response of *disconfirmation*, more specifically, disconfirmation of preperformance standards, results. Because the early work in consumer satisfaction was conducted with predictive expectations as a standard, the phrase "disconfirmation of expectations" or "expectancy disconfirmation" has come to apply to this concept. Many standards consumers bring to the consumption experience can be disconfirmed, so an alternative phrase to describe the discrepancy from a standard could be simply "disconfirmation."

However, because the phrase "disconfirmation" without a valence qualifier is ambiguous as to direction, the phrase "negative disconfirmation" is commonly used to refer to the negative discrepancy that occurs when performance is below standard, and "positive disconfirmation" is used to refer to the positive discrepancy that occurs when performance is above standard. When performance is equal to standards or expectations, a simple confirmation of expectation exists.

Disconfirmation is "generic" with numerous applications. For example, the entire consumption experience can be judged on the degree to which it was better or worse than was expected, as can individual attributes. Other possibilities include dimensions (attribute groups) of performance and the benefits (good aspects) reaped from consumption as separate from the problems (bad aspects) encountered.

Objective (calculated) versus subjective disconfirmation. Early attempts to measure disconfirmation used a discrepancy or "gap" approach. That is, separate survey sections were used to capture, first attribute expectations and, later attribute performance perceptions. Then the performance scores were subtracted from their respective expectation scores and these "gaps" were added. The logic is direct. When attribute performance was higher than its respective expectation, the gap is positive and is considered favorable. Similarly, when expectations were higher, the gap was negative and unfavorable. Satisfaction should increase as the positivity of the gap score increased and it should decrease (contributing to dissatisfaction) with the negativity of the gap score.

Is there evidence for the superiority of subjective disconfirmation (better/worse than) over calculated disconfirmation in the prediction of satisfaction? Yes. A number of studies have examined both the calculated and single-score varieties of disconfirmation, most using rating scale scores. The results of all studies were similar with the majority of the evidence suggesting that the subjective version of disconfirmation correlates more highly with satisfaction scales than do the discrepancy scores. Moreover, when analyzed in an ordering of cause and effect, the following configuration of concepts, as shown in Equation (1), is consistently found to best fit the data:

$$\begin{aligned} &\text{calculated disconfirmation} \\ &\quad \Longrightarrow \text{subjective disconfirmation} \\ &\quad \Longrightarrow \text{satisfaction} \end{aligned} \tag{1}$$

This sequence of events forms the basis for the expectancy disconfirmation model of consumer satisfaction discussed here. As shown, this sequence portrays a calculated expectation–performance discrepancy (if performed) as input to the consumer's subjective interpretation of this difference. The subjective interpretation then becomes the most immediate antecedent of satisfaction. If no "objective" score is available, then a subjective judgment is "sensed." Expectations and performance are implicitly incorporated in the disconfirmation judgment in this sequence.

Operation of disconfirmation in the satisfaction model. Reviews of studies measuring disconfirmation in various forms are now available to suggest that it is a powerful predictor of satisfaction, even when combined with expectation and performance in the manner discussed (Yi, 1990; Szymanski and Henard, 2001; Oliver, 2010). In fact, disconfirmation typically dominates expectation and frequently dominates performance in terms of the strength of effect. There are times, however, when both performance and disconfirmation are input to the same regression that the disconfirmation effect is obscured (becomes nonsignificant). The reason is multicolinearity as disconfirmation is a performance-based concept and both the performance and disconfirmation variables may be highly correlated. When this happens, two regressions must be run separately – the first containing performance and the second containing disconfirmation in its stead.

Short-Term Consequences of Satisfaction

Intention to repurchase. One of the most common results of satisfaction/dissatisfaction is a stated intention (*see* CONSUMER INTENTIONS) to repurchase (or not) in the future. Sometimes, this is posed in surveys in a hypothetical sense as in "if you were in the market for a (generic product), how likely would you buy a (specific brand)?" These scales are ubiquitous in the literature; in particular, it is unusual not to see them being used because the researcher rarely can observe repetitive behavior in a cross-sectional one-shot survey. In commercial research, however, intention scales may not be needed because actual repeat behavior is more easily obtained.

Still another version of intention is the degree to which the consumer splits purchasing between alternatives. Similar to multibrand loyalty, consumers may intend to repatronize one of a set of acceptable alternatives, such as in restaurant dining. In this case, an intention to repatronize a particular establishment is more akin to a probability across choices as opposed to a probability within a choice. It is known, however, that stated intentions without behavioral validation are very unreliable. Consumers frequently overstate their intentions due to a positivity bias in consumer responding. In lieu of other measures of satisfaction validation, however, intention data may be among the best measurement modalities one can achieve.

Complaining/praising. Perhaps the most neglected, infrequently found satisfaction-related concepts in satisfaction surveys are complaining and its polar opposite, praising. This is surprising as the complaining literature is vast and, in fact, was the first of the satisfaction concepts to be extensively studied. Complaining is important because, unlike dissatisfaction, complaining is a behavior. While dissatisfaction and complaining are related, they have been found to be imperfectly correlated. Not all dissatisfied consumers complain so that those who do are very disaffected. Nor are all "complainers" dissatisfied as they may be simply motivated to provide feedback to a firm so that it may improve a current marketplace offering.

In the same vein, the extent of complimenting or praising – the related bipolar concept of complaining – may also be of value to the firm. Praising is not as frequent as complaining, but it does occur. One might view it in the context of an extreme expression of satisfaction, having information value to the firm beyond high performance and satisfaction ratings. This could be particularly important in the service industry where many and varied service providers are involved.

Word of mouth and recommendations. Word of mouth (WOM) is the third post-satisfaction concept discussed here. The nature of recommendations is very closely intertwined with WOM. While WOM can consist of praising or damning (to other consumers as opposed to the firm or its representatives), recommendations are targeted communications to potential purchasers. Note that "recommendations" is a general term and can be either positive (to buy) or negative (not to buy).

Other Satisfaction-Related Comparative Referents

The following are briefly mentioned here because they appear prominently in overall perspectives on the satisfaction response (Oliver, 2010).

Quality. The quality literature (*see* QUALITY FUNCTION DEPLOYMENT (QFD)) predates the satisfaction concept and, in some sense, promoted the emergence of focused attention on satisfaction. In effect, it is a comparison of performance to standards of excellence or perfection as in the four criteria of diamond evaluation (cut, clarity, color, karat weight). The number of citations to this concept is legend, and justice would not be served if any were listed as prototypical.

Equity/inequity. Equity, a comparison of performance to fairness standards, is similarly well-vested. Only a small number of satisfaction studies are available that tackle the core concept (e.g., Oliver and Swan, 1989). Most address unfair pricing (Campbell, 1999), or subdimensions, such as interactional fairness (Blodgett, Hill, and Tax, 1997).

Regret. Regret is of recent vintage, at least in consumer behavior, and is a comparison of performance to "what might have been." Relying heavily on the concept of forgone alternatives, a complete summary can be found in Zeelenberg and Pieters (2007).

Value. This article ends with the concept of value which has taken two forms in the consumer literature. The most common is that of performance compared to sacrifice in acquisition, in effect, a benefit versus cost perspective (Zeithaml, 1988). The second is that of performance against goal attainment, or the second derivative of performance – its intended consequences (Woodruff and Gardial, 1996). The

latter is later subdivided into usage consequences and/or mere possession effects.

There is much more to the study and meaning of satisfaction and more is yet to come, as further discovery on this vast concept continues.

Bibliography

Blodgett, J.G., Hill, D.J., and Tax, S.S. (1997) The effects of distributive, procedural, and interactional justice on postcomplaint behavior. *Journal of Retailing*, **73**, 185–210.

Campbell, M.C. (1999) Perceptions of price unfairness: antecedents and consequences. *Journal of Marketing Research*, **36**, 187–199.

Oliver, R.L. (2006) Customer satisfaction research, in *The Handbook of Marketing Research: Uses, Misuses, and Future Advances* (eds R. Grover and M. Vriens), Sage, Thousand Oaks, pp. 569–587.

Oliver, R.L. (2010) *Satisfaction: A Behavioral Perspective on the Consumer*, 2nd edn, M.E. Sharpe, Armonk, NY.

Oliver, R.L. and Swan, J.E. (1989) Consumer perceptions of interpersonal equity and satisfaction in transactions: a field survey approach. *Journal of Marketing*, **53**, 21–35.

Szymanski, D.M. and Henard, D.H. (2001) Customer satisfaction: a meta-analysis of the empirical evidence. *Journal of the Academy of Marketing Science*, **29**, 16–35.

Woodruff, R.B. and Gardial, S.F. (1996) *Know Your Customer: New Approaches to Understanding Customer Value and Satisfaction*, Blackwell Publishers, Cambridge.

Yi, Y. (1990) A critical review of consumer satisfaction, in *Review of Marketing 1990*, (ed. V.A. Zeithaml.), American Marketing Association, Chicago, pp. 68–123.

Zeelenberg, M. and Pieters, R. (2007) A theory of regret regulation 1.0. *Journal of Consumer Psychology*, **17** (1), 3–18.

Zeithaml, V.A. (1988) Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence. *Journal of Marketing*, **52**, 2–22.

consumer innovativeness

Gerard J. Tellis and Eden Yin

Consumer innovativeness is a construct that deals with how receptive consumers are to new products. Knowledge of consumer innovativeness is crucial for a firm's success given that new products are essential for future growth and success, cost millions of dollars to develop, and their life cycles are becoming increasingly short.

Research on consumer innovativeness is extensive, and yet it has achieved little consensus in definition, measurement, and findings. Consumer innovativeness has been defined as a predisposition or propensity to buy or adopt new products (Midgley and Dowling, 1978; Hirschman, 1980; Steenkamp *et al.*, 1999; Tellis, Yin and Bell, 2009), a willingness to change (Hurt *et al.*, 1977; Im *et al.*, 2003), or a preference for new and different experiences (Hirschman, 1980; Raju, 1980; Venkatraman and Price, 1990).

The measurement of consumer innovativeness has varied greatly across studies partly because of the variation in its definition. The main scales developed to measure this construct are life innovativeness scales and adoptive innovativeness scale. An example of the former is the ability to introduce newness in one's life (Kirton, 1976). Examples of the latter are measuring innovativeness as a tendency to buy new products, which include Raju's scale (1980), Goldsmith and Hofacker's scale (1991), and Baumgartner and Steenkamp's (1996) exploratory product acquisition scale. One problem of these scales is that they have a large number of multiple items that are not substantially different from each other. Thus, the discipline can benefit from a parsimonious measure of consumer innovativeness.

On the basis of a review of the literature, Tellis, Yin and Bell (2009) identify at least ten dimensions of consumer innovativeness: novelty seeking, risk taking, variety seeking, opinion leadership, stimulus variation, habituation, nostalgia, suspicion, effort, and frugality. The first five of these are positively valenced while the latter five are negatively valenced. In a test of these dimensions across 15 countries of the world, Tellis, Yin and Bell (2009) found that the positively valenced measures are prone to social desirability bias and yea saying, while the negatively valenced measures do not suffer from such a bias. Indeed, they found that three or four of the negatively valenced measures could provide a reasonably good estimate of consumer innovativeness across cultures and countries.

Another important measure used to capture consumer innovativeness is a concept called sales takeoff. *Sales takeoff* is the first dramatic increase in the sales of a new product in the early stage of its life cycle. Sales takeoff often marks the beginning of the growth stage of a product's life cycle and can signal its future evolution into a mass market product. Tellis, Stremersch and Yin (2003) and Chandrasekaran and Tellis (2008) use the time to takeoff as a measure for consumer innovativeness across countries. Their finding indicates that clear country clusters of consumer innovativeness emerge based on time to takeoff. For example, in Europe, Scandinavian countries are the most consumer innovative, followed by mid-European countries, and then Mediterranean countries.

Researchers have found some demographic correlates of consumer innovativeness are significant: age, income, education, gender, and mobility. For example, Steenkamp *et al.* (1999) found age to be significant and Steenkamp and Burgess (2002) found income and gender to be significant drivers of consumer innovativeness. In the largest cross-country study on consumer innovativeness, Tellis, Yin and Bell (2009) found age, education, gender, income, and mobility to be significant drivers of consumer innovativeness. Their study is based on data from 5569 consumers across 15 major nations and 9 languages.

Demographics aside, consumer innovativeness is category specific. In particular, women are more eager to buy new home appliances, cosmetics, and food and grocery products, while men are more eager to buy new automobiles and sporting goods. Younger consumers are more eager to buy automobiles than other age groups. Highly educated consumers are more eager to buy financial service (Tellis, Yin and Bell, 2009). Consumer innovativeness also varies across countries and categories. For example, Japanese are most prone to buy new electronics, Brazilians to buy cosmetics, and Swedes to buy

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

food products (Chandrasekaran and Tellis, 2008; Tellis, Yin and Bell, 2009).

See also *choice models; consumer decision making*

Bibliography

Baumgartner, H. and Steenkamp, J.-B.E.M. (1996) Exploratory consumer buying behavior: conceptualization and measurement. *International Journal of Research in Marketing*, **13**, 121–137.

Chandrasekaran, D. and Tellis, G.J. (2008) The global takeoff of new products: culture, wealth, or vanishing differences. *Marketing Science*, **27** (5), 844–860.

Goldsmith, R.E. and Hofacker, C.G. (1991) Measuring consumer innovativeness. *Journal of the Academy of Marketing Science*, **19** (3), 209–222.

Hirschman, E.C. (1980) Innovativeness, novelty seeking, and consumer creativity. *Journal of Consumer Research*, **7** (3), 283–295.

Hurt, H.T., Joseph, K. and Cook, C.D. (1977) Measuring consumer innovativeness. *Journal of the Academy of Marketing Science*, **19** (3), 209–221.

Im, S., Bayus, B.L. and Mason, C.H. (2003) An empirical study of innate consumer innovativeness, personal characteristics and new product adoption behavior. *Journal of the Academy of Marketing Science*, **31** (1), 61–73.

Kirton, M. (1976) Adaptors and innovators: a description and measure. *Journal of Applied Psychology*, **61** (5), 622–629.

Midgley, D.F. and Dowling, G.R. (1978) Innovativeness: the concept and its measurement. *Journal of Consumer Research*, **4**, 229–242.

Raju, P.S. (1980) Optimum stimulation level: its relationship to personality, demographics, and exploratory behavior. *Journal of Consumer Research*, **7**, 272–282.

Steenkamp, J.-B.E.M. and Burgess, S.M. (2002) Optimum stimulation level and exploratory consumer behavior in an emerging consumer market. *International Journal of Research in Marketing*, **19** (2), 131–150.

Steenkamp, J.-B.E.M., Hofstede, F. and Wedel, M. (1999) A cross-national investigation into the individual and national cultural antecedents of consumer innovativeness. *Journal of Marketing*, **63**, 55–69.

Tellis, G., Stremersch, S. and Yin, E. (2003) The international takeoff of new products: the role of economics, culture and country innovativeness. *Marketing Science*, **22** (2), 188–208.

Tellis, G., Yin, E. and Bell, S. (2009) Global consumer innovativeness: cross-country differences and demographic commonalities. *Journal of International Marketing*, **17** (2), 1–22.

Venkatraman, M.P. and Price, L.L. (1990) Differentiating between cognitive and sensory innovativeness. *Journal of Business Research*, **20** (4), 293–315.

social networks

Jacob Goldenberg

A *social network* is usually defined as a graph of nodes, or actors, (e.g., in a marketing context, actors are individual consumers, families, or business entities) that are linked to each other in a meaningful way (e.g., through information exchange or friendship, in the case of individuals, or through trade or collaboration in the case of business entities). These networks are often complex, and social networks can be extremely large in marketing.

It is beyond the scope of this article to note all the works in this field. Several key terms are defined in the box.

In the last two decades, our understanding of social network properties has advanced significantly, and networks have become one of the most attractive fields of research for scholars of almost all social science disciplines, including computer sciences (computer-based networks), biology (e.g., networks of genes), physics, and chemistry. One of the most interesting features of network research is that researchers from diverse fields typically work on similar problems, use similar definitions, and they obtain consistent results concerning networks dynamics (as a result, a typical network paper includes references from multiple fields). Some examples of well-known network properties found to be common to many networks (briefly defined in the box) include a *high clustering coefficient* (Watts, 1999), *short distances or paths* (see Newman, 2001; Watts and Strogatz, 1998), a *scale-free degree distribution* (Albert and Barabasi, 2002; Amaral *et al*., 2000; Barabási and Albert, 1999; Jeong *et al*., 2000), and a common *betweenness centrality distribution* (Chen *et al*., 2007; Kitsak *et al*., 2007).

Social network literature can be classified into two research themes: (i) network formation (see among many others: Albert and Barabasi, 2002; Barabási and Albert, 1999; Davidsen, Ebel, and Bornholdt, 2002; Deroïan, 2006; Kryssanov *et al*., 2008; Kryssanov *et al*., 2006; Watts, 1999; Zhang and Liu, 2006; Zhou and Mondragon, 2004) and (ii) information dissemination over networks and network connectivity (see among others: Brown, 1981; Goldenberg, Libai, and Muller, 2001; Goldenberg *et al*., 2005; Kocsis and Kun, 2008; Rogers, Ascroft, and Rolling, 1970; Valente, 1995; Young, 2006).

In the first research stream, recent efforts have been applied to understand how network structures and properties emerge, and to identify the local rules of behavior that generate the formation of these structures. One topic of study in this stream is scale-free emergence. Barabási and Albert 1999 mapped the topology of a portion of the Web, and reported two findings: some nodes (termed *hubs*) had an exceptionally greater number of links than other nodes, and the number of links connecting to a single node followed a power-law distribution. After discovering that several other networks, including social and biological networks, typically had heavy-tailed degree distributions, they called this structure a *scale-free network*. Amaral *et al*. demonstrated that most real-world networks can be classified into two broad categories according to the slope of the distribution (usually plotted on a log–log scale). Barabási and Albert proposed the *preferential attachment* mechanism to explain the appearance of this power-law distribution. This mechanism (with many variations) essentially assumes that any new node joining the network has a higher probability to link to a node with a higher degree. Alternative mechanisms, such as transitivity (in which a friend of my friend is likely to become my friend), have also been since suggested.

Marketing literature is more strongly oriented to the second research stream (for a recommended review, see Van den Bulte and Wuyts, 2007), with a recent focus on information dissemination, accepted as the core process driving innovation adoption (see Godes *et al*., 2005; Goldenberg *et al*., 2005; Hogan, Lemon, and Libai, 2005; Rogers, 2003; Trusov, Bodapati, and Bucklin, 2008; Valente, 1995; Van den Bulte and Joshi, 2007; Van den Bulte and Wuyts, 2007; Watts and Dodds, 2007). For example, a special session on networks was held at a recent Marketing Science conference, to discuss networks and related issues such as network population sampling (Ebbes *et al*., 2008), the role of spatial proximity in diffusion (Barrot *et al*., 2008), the economic value of social interactions (Stephen and Toubia, 2010), and a study of pharmaceutical social

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

networks (Iyengar, Valente, and van Den Bulte, 2008). Research has also empirically established the intuitive understanding that innovations propagate through peer recommendations and word of mouth, which are strongly dependent on the structure of the social network. Several additional interesting issues are presented briefly below.

The Strength of Weak Ties

Granovetter (1973) found that weak ties are sometimes more important than strong ties in information seeking and innovation dissemination. Strongly tied individuals form cliques that have a tendency to conformity. Owing to their fundamental similarity, any single clique member also knows more or less what the other members know. Therefore, to discover new information or gain new insights, clique members must look beyond the clique to other, more distant friends, and acquaintances. This weak-ties strength was found to have significant implications for consumer behavior issues.

Small Worlds and Degrees of Separation

The scale-free structure is one explanation for evidence on the nature of the dissemination process and the role of high connectivity in this process: The hubs receive information first and spread it through the network. Another explanation is the small world structure and the high clustering coefficient (Watts and Strogatz, 1998).

The small world theory is based on the premise that any arbitrary pair of nodes in the network is linked by a genuinely short chain of individuals (nodes). The 1967 small world experiment by Milgram, 1967 gave rise to the famous phrase "six degrees of separation." In this experiment, a sample of individuals was asked to reach a specific target person by passing a message along a chain of acquaintances (all pairs were in the United States). Successful chains (which reached their destination) turned out to be approximately six steps long, which is remarkably short. Note, however, that this finding may be biased by the fact that the majority of chains in that study actually failed to complete. Despite claims of the experiment's questionable methodology, it is widely accepted that a social network requires a small number of steps to connect all its members. A recent electronic small world experiment found that approximately five to seven degrees of separation are sufficient for connecting any two people through e-mail (Watts, 1999), although the majority of chains in this experiment were also not completed.

According to the small world model, most nodes are not neighbors of one another, but can be reached from any other node using only a small number of jumps. A small world network, where nodes represent people and edges (or links) that connect acquaintances, captures the phenomenon of strangers who become linked through a mutual acquaintance. Watts and Strogatz (1998) noted that networks can be classified according to two independent structural features: the clustering coefficient and shortest path (distance). Random networks, for example, exhibit a small shortest pathlength and a small clustering coefficient. Indeed, many real-world networks not only have a small average shortest pathlength but also a clustering coefficient significantly higher than expected by random chance. This alone explains the rapid dissemination of information that occurs over a network.

The Role of Key Individuals

Research suggesting that a relatively small number of people have substantial influence on the opinions and decisions of the majority can be traced back at least 50 years to the seminal work by Katz and Lazarsfeld (1955). Broadly speaking, influential people are thought to have three important traits: (i) they are persuasive (perhaps even charismatic); (ii) they know a lot (i.e., are experts); and (iii) they have large number of social ties, i.e., they know a lot of people.

Hyper-influential individuals (also known as *opinion leaders*, *mavens*, or *hubs*) are one class of consumers that has always attracted the attention of researchers and practitioners. Extensive studies of the behavior of agents in social systems have been conducted at the social and psychological level (Reingen and Kernan, 1986). While different terms for these agents are typically used interchangeably, the constructs are not identical. Opinion leaders

and mavens (Coulter, Larence, and Price, 2002; Feick and Price, 1987) are thought to have expertise in their area of influence, while hubs are individuals whose distinctive feature is their large number of social ties.

The literature on opinion leaders is relatively broad and opinion leaders have been studied in a variety of areas including marketing, public opinion, health care, communication, education, agriculture, and epidemiology. Until recently, researchers have broadly concurred that opinion leaders can have a major impact on opinion formation and change, and that a small group of influential opinion leaders may accelerate or block the adoption of a product in an entire market. Since 2008, this point has triggered heated debate. Some argue that influentials (or more precisely, individuals who represent nodes with an exceptionally large degree) have only a marginal influence on adoption speed (e.g., Watts and Dodds, 2007), yet according to counterarguments, influentials adopt sooner than other people. This occurs not because they are innovative but rather because influentials are exposed earlier to an innovation due to their numerous social links (e.g., Goldenberg *et al.*, 2009). As a result, highly connected individuals have a significant impact on the overall adoption speed of an innovation and in some cases even on the ultimate market size.

Social network metrics

Betweenness: Betweenness reflects the extent to which a node connects other pairs of nodes in the network (through the shortest path).

Centrality: Centrality reflects the social "power" of a node, based on how well it connects the network. The more common centrality measures are *degree*, *betweenness*, and *closeness*.

Closeness: Closeness reflects the proximity of all other individuals in a network (directly or indirectly) to the node. Closeness is the inverse of the sum of the shortest distances between each node and every other node in the network.

Clustering coefficient: The clustering coefficient measures the likelihood that two acquaintances of a node are acquaintances themselves: if A knows both B and C, a high coefficient in this network implies that the probability that B knows C is high.

Degree: The degree is the number of ties connected to each node.

Density: Network density is the proportion of ties in a network relative to the total number of possible ties.

Pathlength: Pathlength is the distance between pairs of nodes in the network.

Scale-free network: Scale-free networks have a degree distribution that is easy to view as a power law: $n(\text{degree} = x) = -ax^y$.

Structural cohesion: Structural cohesion is measured by the minimum number of nodes required to dismantle a group of nodes.

Structural equivalence: Structural equivalence is the extent to which nodes share a common set of ties.

Bibliography

Albert, R. and Barabasi, A.L. (2002) Statistical mechanics of complex networks, *Reviews of Modern Physics*, **74** (1), 47–97.

Amaral, L.A.N., Scala, A., Barthélémy, M., and Stanley, H.E. (2000) Classes of small-world networks, *Proceedings of the National Academy of Sciences of the United States of America*, **97** (21), 11149–11152.

Barabási, A.L. and Albert, R. (1999) Emergence of scaling in random networks,. *Science*, **286** (5439), 509–512.

Barrot, C., Rangaswamy, A., Albers, S., and Shaikh, N.I. (2008) The Role of Spatial Proximity in the Adoption of a Digital Product, Working Paper.

Brown, L.A. (1981) *Innovation Diffusion: A New Perspective*, Methuen, London and New York.

Chen, Y.P., Paul, G., Cohen, R. *et al.* (2007) Percolation theory applied to measures of fragmentation in social networks. *Physical Review E*, **75** (4), 1–7.

Coulter, R.A., Larence, F.F., and Price, L.L. (2002) Changing faces: cosmetics opinion leadership among women in the new Hungary. *European Journal of Marketing*, **36** (11), 1287–1308.

Davidsen, J., Ebel, H., and Bornholdt, S. (2002) Emergence of a small world from local interactions: modeling acquaintance networks. *Physical Review Letters*, **88** (12), 1–4.

Deroïan, F. (2006) Formation of a communication network under perfect foresight. *Theory and Decision*, **61** (3), 191–204.

Ebbes, P., Zan, H., Rangaswamy, A., and Thadakamall, H.P. (2008) Sampling Large-Scale Social Networks: The Good, the Bad, and the Ugly, Working Paper.

Feick, L.F. and Price, L.L. (1987) The market maven: a diffuser of marketplace information. *Journal of Marketing*, **51** (1), 83–97.

Godes, D., Mayzlin, D., Chen, Y. *et al*. (2005) The firm's management of social interactions. *Marketing Letters*, **16** (3), 415–428.

Goldenberg, J., Libai, B., and Muller, E. (2001) Talk of the network: a complex systems look at the underlying process of word-of-mouth. *Marketing Letters*, **12** (3), 211–223.

Goldenberg, J., Shavitt, Y., Shir, E., and Solomon, S. (2005) Distributive immunization of networks against viruses using the 'Honey Pots' architecture. *Nature Physics*, **1**, 184–188.

Goldenberg, J., Sangman, H., Lehmann, D.R., and Weon-Hong, J.W. (2009) *The Role of Hubs in the Adoption Processes*, Hebrew university of Jerusalem, Jerusalem, Israel. working paper.

Granovetter, M.S. (1973) The strength of weak ties. *American Journal of Sociology*, **78** (6), 1360–1380.

Hogan, J.E., Lemon, K.N., and Libai, B. (2005) Quantifying the ripple: word-of-mouth and advertising effectiveness. *Journal of Advertising Research*, **44** (03), 271–280.

Iyengar, R., Valente, T., and van Den Bulte, C. (2008) *Opinion Leadership and Social Contagion in New Product Diffusion*, University of Pennsylvania. working paper

Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., and Barabasi, A.L. (2000) The large-scale organization of metabolic networks. *Nature*, **407** (6804), 651.

Katz, E. and Lazarsfeld, P.F. (1955) *Personal Influence; The Part Played by People in the Flow of Mass Communications. Glencoe, Ill*, Columbia University. Bureau of Applied Social, Research, Free Press.

Kitsak, M., Havlin, S., Paul, G. *et al*. (2007) Betweenness centrality of fractal and nonfractal scale-free model networks and tests on real networks. *Physical Review E*, **75** (5), 1–8.

Kocsis, G. and Kun, F. (2008) The effect of network topologies on the spreading of technological developments. *Journal of Statistical Mechanics-Theory and Experiment*, **P10014**, 1–13.

Kryssanov, V.V., Rinaldo, F.J., Kuleshov, E.L., and Ogawa, H. (2006) Modeling the Dynamics of Social Networks. *ICE-B*, 242–249.

Kryssanov, V.V., Rinaldo, F.J., Kuleshov, E.L., and Ogawa, H. (2008) A Hidden Variable Approach to Analyze "Hidden" Dynamics of Social Networks, in Why Context Matters.

Milgram, S. (1967) The small world problem. *Psychology today*, **2** (1), 60–67.

Newman, M.E.J. (2001) The structure and function of complex networks. *Structure*, **45** (2), 167–256.

Reingen, P.H. and Kernan, J.B. (1986) Analysis of referral networks in marketing: methods and illustration. *Journal of Marketing Research*, **23** (4), 370–378.

Rogers, E.M. (2003) *Attributes of Innovations and their Rate of Adoption and Innovativeness and Adopter Categories. Diffusion of Innovations*, Free Press, New York.

Rogers, E.M., Ascroft, J., and Rolling, N.G. (1970) *As Reported in Valente, Network Models of the Diffusion of Innovations*, Hampton Press, Cresskill.

Stephen, A.T. and Toubia, O. (2010) Deriving value from social commerce networks. *Journal of Marketing Research*, **47** (2), 215–228.

Trusov, M., Bodapati, A.V., and Bucklin, R.E. (2008) Determining Influential Users in Internet Social Networks, working paper.

Valente, T.W. (1995) *Network Models of the Diffusion of Innovations*, Hampton Press, Cresskill.

Van den Bulte, C. and Joshi, Y.V. (2007) New product diffusion with influentials and imitators. *Marketing Science*, **26** (3), 400–421.

Van den Bulte, C. and Wuyts, S. (2007) *Social Networks and Marketing*, Marketing Science Institute.

Watts, D.J. (1999) *Small Worlds: the Dynamics of Networks Between Order and Randomness*, Princeton University Press, Princeton.

Watts, D.J. and Dodds, P.S. (2007) Influentials, networks, and public opinion formation. *Journal of Consumer Research*, **34** (4), 441.

Watts, D.J. and Strogatz, S.H. (1998) Collective dynamics of 'small-world' networks. *Nature*, **393** (6684), 440–442.

Young, H.P. (2006) The Diffusion of Innovations in Social Networks, Economy as an Evolving Complex System, III, pp. 267–281

Zhang, S. and Liu, J. (2006) *From Local Behaviors to the Dynamics in an Agent Network*, Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence: IEEE Computer Society, Hong Kong, China, 572–580.

Zhou, S. and Mondragon, R.J. (2004) The rich-club phenomenon in the Internet topology. *IEEE Communications Letters*, **8** (3), 180–182.

persuasion

Michal Herzenstein

Persuasion is an active attempt to change beliefs, attitudes, preferences, or behaviors. The extent of persuasion is measured by changes to the above (measured before and after the persuasion attempt). This article focuses on various approaches to persuasion.

Cognitive Approaches

Persuasion will take place if cognitive processing, conscious or unconscious, has occurred. An example for conscious cognitive approaches is comparative judgment theory, which suggests that nothing is judged in isolation and everything can be seen as good or bad depending on what it is compared with (framing and prospect theory, Tversky, and Kahneman, 1979; contrast and assimilation effects, Hovland, Harvey, and Sherif, 1957). Examples for unconscious cognitive processes are implicit memory (Schacter, 1987) and mere exposure theories (Zajonc, 1968) that show that even under conditions of limited processing resources, exposure can influence judgment (albeit without awareness).

Affective Approaches

People are persuaded because the communication evokes certain affects. Both positive affects (such as humor, love, or pride) and negative affects (such as fear, disturbance, or unpleasantness) are commonly induced. Persuasion occurs because people try to maintain their positive moods or avoid negative moods. One important theory that explains how to use emotions in persuasion is classical conditioning. On the basis of Pavlov's experiments with dogs, this theory suggests that pairing brands with positive unconditioned stimuli (familiar stimuli that automatically produce affective responses) will elicit the desired affective response toward the brand itself (rather than the stimuli). For example, a brand is paired with a familiar and loved tune in a commercial (Gorn, 1982). The tune elicits positive emotions that are paired with the brand. Afterwards, when the brand is encountered without the tune (i.e., in the store) it elicits the same emotions.

Motivational Approaches

Persuasion happens when people wish to be persuaded. Subtle motivational techniques are preferred over obvious persuasion attempts (Kardes, 2001). The most relevant principle is consistency – people like consistency and dislike inconsistency. They dislike it so much that they are willing to change beliefs, attitudes, and behaviors in order to eliminate the inconsistency. A famous example of this principle is cognitive dissonance (Festinger, 1957), which suggests that people strive for attitude–behavior consistency. When attitudes are inconsistent with past behaviors (dissonance), an unpleasant tension is produced, which motivates people to reduce that dissonance. The result is a shift in attitudes that increases attitude–behavior consistency.

Self-Persuasion

Persuasion that results from elaboration on the true merit of the brand leads to strong accessible attitudes, and therefore is preferred over persuasion that results from peripheral cues such as music or humor. Self-persuasion approaches try to understand how persuasive communications induce people to think more extensively about the brand or advocated position. One of the most useful models is the elaboration likelihood model of persuasion (ELM) (Petty and Cacioppo, 1981). ELM proposes two routes to persuasion: the central route – requires effortful processing of the message; and the peripheral route – requires minimal thinking. The central route leads to strong accessible attitudes, but it will be taken only when the recipient is both motivated and able to elaborate on the message.

Social Influence

Communicators who wish to move people in their direction should appeal to a limited set of deeply rooted human drives and needs. According to Cialdini (2001) there are six principles that lead to persuasion. (i) Liking – sellers should like their customers. People like people who like them, and they are more likely to say yes to them because they feel they are in good hands. (ii) Reciprocity – people want to give back to those who have given to them.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Companies who give their clients something meaningful, tailored, and unexpected are often repaid. (iii) Social proof – people like to follow the crowd. Use the many to persuade the few. (iv) Consistency – when people make a public commitment, they almost always back those words with actions. (v) Authority – people like to follow the lead of legitimate experts. (vi) Scarcity – people are afraid to miss out or lose by not choosing something scarce.

In conclusion, persuasion can be attained in various ways, and choosing the proper appeal is crucial. Any communicator who wishes to move people in his or her direction should exercise some "detective work" and seek out the right time, and the right place, to put the right approach to use with his/her target audience (Cialdini, 2001).

See also *attitudes; emotion; implicit consumer cognition; social influence*

Bibliography

Cialdini, R.B. (2001) Harnessing the science of persuasion. *Harvard Business Review*, **79** (9), 72–79.

Festinger, L. (1957) *A Theory of Cognitive Dissonance*, Row and Peterson, Evanston.

Gorn, G.J. (1982) The effect of music in advertising on choice behavior: a classical conditioning approach. *Journal of Marketing*, **46**, 94–101.

Hovland, C.I., Harvey, O.J., and Sherif, M. (1957) Assimilation and contrast effects in reactions to communication and attitude change. *The Journal of Abnormal and Social Psychology*, **55** (2), 244–252.

Kardes, F.R. (2001) *Consumer Behavior and Managerial Decision Making*, Pearson Education, Inc., Upper Saddle River.

Petty, R.E. and Cacioppo, J.T. (1981) *Attitudes and Persuasion: Classic and Contemporary Approaches*, William C. Brown, Dubuque.

Schacter, D.L. (1987) Implicit memory: history and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **13**, 501–518.

Tversky, A. and Kahneman, D. (1979) Prospect theory: an analysis of decision under risk. *Econometrica*, **47** (2), 263–291.

Zajonc, R.B. (1968) Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, **9** (2), 1–27.

consumer involvement

Judith Lynne Zaichkowsky

Involvement is a motivational variable in consumer behavior and can be defined as *A person's perceived relevance of the object based on inherent needs, values and interests* (Zaichkowsky, 1985, 1986). Involvement is used to describe the level of consumer interest, search, or complex decision making toward an object. The object of involvement may be a product, a service, a situation, or an advertisement. Low involvement implies inertia, and high involvement implies a great deal of activation and it is found to influence consumer decision making and interactive communications. A consumer's level of involvement can be used as a segmenting variable to further target the market by marketing managers.

The history of the term *involvement* in the marketing literature has two major roots: one in the advertising research literature (Krugman, 1965, 1967) and the other in the consumer behavior literature (Howard and Sheth, 1969). In applying learning theory to TV copy testing, Krugman 1965 found that when a series of ads were presented, those presented first and last were better remembered than those ads in the middle of the series. This finding showed the same primacy and recency effects found by Hovland 1957 in the learning of non-ego involving material. Krugman hypothesized what advertising and non-ego involving material had in common were low levels of involvement, and operationalized it as the number of "bridging experiences, connections or personal references" per minute that the viewer made between his own life and the advertisement. Therefore personal involvement impacted on response to advertising.

Consumer involvement in the marketing and consumer behavior literature grew out of the realization that a great deal of consumer behavior does not involve extensive search for information or a comprehensive evaluation of choice alternatives (Olshavsky and Granbois, 1979). The consumer makes dozens of mundane decisions and choices each day and it is inappropriate to assume consumers actively process and think about each decision (Kassarjian, 1978). In this domain, many studies looked at the act of purchase, in addition to the product category itself, and determined the situations of purchase were highly influential in determining the level of consumer decision making activity (Clarke and Belk, 1978).

Because involvement was discussed in terms of products, advertisements, and purchase situations, early quantitative studies used various manipulations or single item measures and outcomes to represent and capture the concept. In a response to the plethora of definitions and measures, Zaichkowsky (1985) proposed a single definition and semantic differential scale to be used across all domains to capture the abstract concept of involvement on one dimension. Laurent and Kapferer (1985), on the other hand, proposed a measurement of product involvement which captured five facets or dimensions of product involvement: personal interest, importance of negative consequences, subjective probability of mispurchase, pleasure value, and sign value. These two studies represented the beginning of a long debate on how best to measure involvement.

It was clear that involvement was not only cognitive, but could also be emotional in nature (Petty, Cacioppo, and Schumann, 1983). The original Zaichkowsky Personal Involvement Inventory (PII) was later reduced to 10 items (Zaichkowsky, 1994) which better represented a balance of cognitive and emotional involvement items. This revised PII fitted in well with the idea that there are four quadrants in consumer behavior research: low and high involvement on one axis, and emotional and cognitive involvement on the other (Vaughn, 1988).

In recent years, the research literature indicates the concept of consumer involvement is widely embraced across a multiple of disciplines from agriculture (e.g., Verbeke and Vackier, 2004) to information systems (e.g., Koufaris, 2002). The role of involvement in consumer behavior has been so well accepted that it is completely incorporated in the field and no textbook in consumer behavior seems complete without a discussion of the topic.

Bibliography

Clarke, K. and Belk, R. (1978) The effects of product involvement and task definition on anticipated

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

consumer effort, in *Advances in Consumer Research*, vol. **5** (ed. H.K. Hunt), Association for Consumer Research, Ann Arbour, pp. 313–318.

Hovland, C.I. (1957) *The Order of Presentation in Persuasion*, Yale University Press, New Haven.

Howard, J.A. and Sheth, J.N. (1969) *The Theory of Buyer Behavior*, John Wiley & Sons, Inc., New York.

Kassarjian, H.H. (1978) Presidential address, in *Advances in Consumer Research*, vol. **5** (ed. H.K.Hunt), Association for Consumer Research, Ann Arbour, pp. 31–34.

Koufaris, M. (2002) Applying the technology acceptance model and flow theory to online consumer behavior. *Information Systems Research*, **13**, 205–223.

Krugman, H.E. (1965) The impact of television advertising involvement. *Public Opinion Quarterly*, **29**, 349–356.

Krugman, H.E. (1967) The measurement of advertising involvement. *Public Opinion Quarterly*, **30**, 583–596.

Laurent, G. and Kapferer, J.N. (1985) Measuring consumer involvement profiles. *Journal of Marketing Research*, **22**, 41–53.

Olshavsky, R. and Granbois, D.H. (1979) Consumer decision making-fact or fiction? *Journal of Consumer Research*, **6**, 93–100.

Petty, R.E., Cacioppo, J.T., and Schumann, D. (1983) Central and peripheral routes to advertising effectiveness: the moderating role of involvement. *Journal of Consumer Research*, **10**, 135–146.

Vaughn, R. (1988) How advertising works: a planning model revisited. *Journal of Advertising Research*, **26** (1), 57–66.

Verbeke, W. and Vackier, I. (2004) Profile and effects of consumer involvement in fresh meat. *Meat Science*, **67**, 159–168.

Zaichkowsky, J.L. (1985) Measuring the involvement construct. *Journal of Consumer Research*, **12**, 341–352.

Zaichkowsky, J.L. (1986) Conceptualizing involvement. *Journal of Advertising*, **15** (2), 4–14.

Zaichkowsky, J.L. (1994) The personal involvement inventory: reduction, revision, and application to advertising. *Journal of Advertising*, **23** (4), 59–70.

consumer desire

Richard P. Bagozzi

As a scientific concept, desire is relatively new in consumer research. Two construals of desires have been proffered by consumer researchers. One is championed by Belk and colleagues (e.g., Belk, Ger, and Askegaard, 1997, 2003). These authors conceive of desires as "belief-based passions that involve longing, yearning, and fervently wishing for something" and can be expressed in metaphors for "hunger (or thirst), sexual lust, and addiction" (Belk, Ger, and Askegaard, 1997, p. 24). Belk, Ger, and Askegaard (2003) characterize desire as "a powerful emotion" (p. 343) observed in a state of tension between feelings of seduction and morality (p. 345) and distinguish it from wants and needs (p. 328). While derived from grounded research with consumers, this conceptualization of desires seems overly narrow, limited too much to biological kinds of desire, and exclusionary of less intense forms of everyday desire functioning in consumer behavior.

A second conceptualization of desire, while overlapping with that formulated by Belk, Ger, and Askegaard (2003), comes from the subfields of philosophy known as the *philosophy of mind* and the *philosophy of action* (e.g., Bishop, 1989; Davis, 1997; Mele, 2003). Davis (1997) distinguishes between volitive and appetitive desires. A volitive desire is "synonymous with *want*, *wish*, and *would like*, and appears as a *transitive verb* in sentences like 'I desire to . . . ' and 'I desire . . . '" (Davis, 1997, p. 136, emphasis in original). For example, "John would like to apply to Harvard" and "Mary wants intellectual stimulation in a movie" are volitive desires. By contrast, an appetitive desire has "the near synonyms *appetite*, *hungering*, *craving*, *yearning*, *longing*, and *urge*, and appears as a *noun* in sentences like 'I have a desire to . . . ' and 'I have a desire for . . . ,' [moreover] objects of appetitive desire are *appealing*, things we *view with pleasure*" (Davis, 1997, p. 136, emphasis in original). For instance, "Silvia has a longing for her birthplace" and "Paul has a craving for sushi" are appetitive desires. Davis points out that volitive and appetitive desires are logically independent and can exist empirically in distinct ways:

> We often want to eat, for social or nutritional reasons, when we have no appetite and view the prospect of eating without pleasure. We desire to eat, but have no desire to. On the other hand we may have a ravenous appetite and find the prospect of eating terribly appealing and yet not want to eat because we are on a diet. (1997, p. 136)

Bagozzi (1992, pp. 183–194) proposed that desires are fundamental psychological events or states (distinct from cognitive, evaluative, and affective reasons for acting) that convert reasons for acting into intentions to act. Such desires have been termed *behavioral desires* or *action desires*. Mele (1995) calls such desires *extrinsic desires* (i.e., desiring to act as a means to an end). Desires also exist as intrinsic desires (i.e., desiring something for its own sake or as an end; Mele, 1995). An important form of intrinsic desires is goal desires.

Desires perform three functions. First, they motivate our goal intentions and our behavioral or action intentions (*see* CONSUMER INTENTIONS). One way they do this is automatically and nonconsciously or nondeliberatively through what Damasio (1994, pp. 173–174) called the *somatic-marker hypothesis*. That is, prior to conscious processing of pros and cons characteristic of rational decision making, people experience pleasant or unpleasant feelings that highlight options and create either positive or negative biases, which favor or eliminate options from consideration. Such unconscious processes influence or bias a number of antecedents to decision making and indeed can form the basis for certain desires. It is likely that declarative knowledge processed rationally by consumers (with regard to facts, alternative goals and brands, consequences of consumption, and various expectations one has) is influenced by unconscious preference biases residing in the brain and arising from previous emotional experience associated with similar decision problems. Bechara *et al.* (1997) present research showing that such covert processes bias decision making prior to cognitive evaluation and reasoning and without awareness occurring

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

on the part of decision makers. The authors suggested that the unconscious processes guide or shape behavior, before conscious processing commences, and function to produce better decisions, especially to the extent that learning accumulates as a consequence of previous rewards and punishments, which become stored as nondeclarative dispositional knowledge. It is possible that some desires, especially appetitive ones, develop in this way and become the basis for goal desires and even some behavioral desires.

Now consider a second function of desires. People are often aware of their desires, and desires seem to function consciously in many decision-making settings. This is especially true for volitive desires. It is common for a decision maker to have many reasons for action, some of which might even conflict or constitute reasons for not acting in the very same decision context. In such cases, desires serve to integrate or summarize a decision maker's overall felt urge to act as a function of multiple, mixed reasons for action. Reasons for action are appraised, combined, and transformed into a motivation to act. Whether this happens as a deterministic resultant of competing forces in response to reasons for action, or follows learned rules of weighting and consolidation, remains to be studied. But in the face of multiple reasons for acting and not acting, in which some reasons are determinative whereas others are not, people subjectively experience and express their final felt urge to act.

A third function for desires is to induce an intention; a goal desire incites a goal intention, and a behavioral desire evokes an implementation intention. (*see* CONSUMER INTENTIONS). Desires harbor energy in an action-tendency or consummatory sense, but without precise direction and without a personal commitment to act. Intentions provide precise direction and personal commitment in this sense. These are the primary differences between desires and intentions. Yet another distinction is that intentions, but not desires, can entail a plan to act, though not all intentions necessarily contain or imply plans. Perugini and Bagozzi (2004) explored additional distinctions between desires and intentions.

Left unchecked, desires tend to influence intentions. This happens for goal intentions as well as for action or behavioral intentions (*see* CONSUMER INTENTIONS). An important issue for research is how desires are self-regulated (*see* Bagozzi, 2006 and SELF-REGULATION). As we becomes aware of our desires, we may reflect upon them and apply self-evaluative standards. In this regard we may ask ourselves (figuratively or literally) whether we are the sort of person who should have, and act, on these desires. The answer may encompass a decision to cancel, override, or postpone implementation of the desire through intention formation. The decision might be influenced by one's social, moral, or self-conscious emotions (e.g., pride, gratitude, guilt, shame, embarrassment, contempt, disgust, anger), by one's ethics, moral beliefs, or character, by one's personal or social identity, or by one's feelings of empathy, love, or affection for another person implicated by the desire one way or the other. In parallel way, a person might come to question why he/she has no desire for a goal or to act and then embrace such a desire as a function of determinants similar to those mentioned in the aforementioned sentence.

Bibliography

Bagozzi, R.P. (1992) The self-regulation of attitudes, intentions, and behavior. *Social Psychology Quarterly*, **55**, 178–204.

Bagozzi, R.P. (2006) Explaining consumer behavior and consumer action: from fragmentation to unity. *Seoul Journal of Business*, **12**, 111–143.

Bechara, A., Damasio, H., Tranel, D., and Damasio, A. (1997) Deciding advantageously before knowing the advantageous strategy. *Science*, **275**, 1293–1295.

Belk, R.W., Ger, G., and Askegaard, S. (1997) Consumer desire in three cultures: results from projective research, in *Advances in Consumer Research*, vol. **24** (eds M. Brucks and D.J. MacInnis) Association for Consumer Research, Provo, pp. 24–28.

Belk, R.W., Ger, G., and Askegaard, S. (2003) The fire of desire: a multisided inquiry into consumer passion. *Journal of Consumer Research*, **30**, 326–351.

Bishop, J. (1989) *Natural Agency: An Essay on the Causal Theory of Action*, Cambridge University Press, Cambridge.

Damasio, A.R. (1994) *Descartes' Error: Emotion, Reason, and the Human Brain*, Avon Books, New York.

Davis, W.A. (1997) A causal theory of intending, in *Philosophy of Action* (ed. A.R. Mele), Oxford University Press, Oxford, pp. 131–148.

Mele, A.R. (1995) Motivation: essentially motivation-constituting attitudes. *Philosophical Review*, **104**, 387–423.

Mele, A.R. (2003) *Motivation and Agency*, Oxford University Press, Oxford.

Perugini, M. and Bagozzi, R.P. (2004) The distinction between desires and intentions. *European Journal of Social Psychology*, **34**, 69–84.

opinion leadership and market mavens

Ronald Earl Goldsmith

Social Communication and Consumer Behavior

Consumers use information to make shopping and buying decisions. They get some of this information from marketer-dominated sources such as advertisements, salespersons, brochures, packages, web pages, and promotions. Another and perhaps more important source of information comes in the form of social communication (*see* SOCIAL INFLUENCE). This is the influence consumers have on other consumers. Also termed *non-marketer-dominated* sources, the influence of other consumers can be passive, where consumers simply observe others and imitate them, or active, where consumers talk to each other, seeking or offering advice and information as part of everyday life (Weimann, 1994).

This final form of social communication is often termed *word of mouth*.Word of mouth can be incidental, where topics covering products, advertisements, stores, brands, and buying can be part of ordinary conversations. Consumers can also actively seek advice and information from other consumers or they can actively give information. The former describes opinion seeking and the latter opinion leadership. Opinion leaders are consumers who actively transmit information to other consumers either on their own initiative or after information seekers solicit them. Some consumers called *market mavens* are especially involved in the marketplace (*see* CONSUMER INVOLVEMENT). They are knowledgeable and actively influence other consumers. Because opinion leaders and market mavens exert such a powerful influence on sales, marketers are keen to identify them and to persuade them to promote their brands.

Opinion Leadership

Most opinion leadership takes the form of product category or domain-specific opinion leadership where consumers spread word of mouth about specific types of products, such as clothing, movies, cars, food, and so on (Flynn, Goldsmith, and Eastman, 1996). The chief antecedent or motivator for opinion leadership is likely involvement, or the interest, enthusiasm, and excitement consumers feel for their favorite categories. Opinion leaders are not only involved in their favorite domains, they are knowledgeable about them, widely exposed to marketer-dominated sources of information, eager to buy new products in the domain (*see* CONSUMER INNOVATIVENESS), and motivated to influence others. Opinion leadership occurs both off-line in the physical world as well as online in cyberspace, especially in SOCIAL NETWORKS where a small number of enthusiastic consumers dominate the flow of information across the network (Iyengar, Van den Bulte, and Valente, 2008).

In addition, some consumers act as general opinion leaders where their influence cuts across several domains (Keller and Berry, 2003). This general type of opinion leadership suggests that some consumers are especially important in the spread of information through social connections. The most thorough examination of this general marketplace behavior occurs in the research stream devoted to market mavenism.

Market Mavens

Market mavens were first described by Feick and Price (1987, p. 85) as: "individuals who have information about many kinds of products, places to shop, and other facets of markets, and initiate discussion with consumers and respond to request from consumers from market information." Since then, researchers have extended the understanding of this phenomenon. Market mavens cannot be especially defined by their demographics, but do thrive in different countries around the world (Chelminski and Coulter, 2007). These consumers pay close attention to ads and brands, they like to shop, they think about shopping and buying, they like to talk to others about ads, stores, products, shopping, and consuming in general. They clip and trade coupons. They know the best places to buy and actively recommend them to their friends. They are aware of and eager to buy new products. They act as general opinion leaders but differ from them through their greater engagement in the marketplace and variety of market-related behaviors.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Bibliography

Chelminski, P. and Coulter, R.A. (2007) On market mavens and consumer self-confidence: a cross-cultural study. *Psychology and Marketing*, **24** (1), 69–91.

Feick, L.F. and Price, L.L. (1987) The market maven: a diffuser of marketplace information. *Journal of Marketing*, **51**, 83–97.

Flynn, L.R., Goldsmith, R.E., and Eastman, J.K. (1996) Opinion leaders and opinion seekers: two new measurement scales. *Journal of the Academy of Marketing Science*, **24** (2), 137–147.

Goldsmith, R.E., Flynn, L.R., and Goldsmith, E.B. (2003) Innovative consumers and market mavens. *Journal of Marketing Theory and Practice*, **11** (4), 54–65.

Iyengar, R., Van den Bulte, C., and Valente, T.W. (2008) Opinion Leadership and Social Contagion in New Product Diffusion. Report No: 08-120, Marketing Science Institute, Cambridge.

Keller, E.B. and Berry, J.L. (2003) *The Influentials: One American in Ten Tells the Other Nine How to Vote, Where to Eat, and What to Buy*, Free Press, New York.

Weimann, G. (1994) *The Influentials*, State University of New York Press, Albany.

optimum stimulation level

Jan-Benedict E. M. Steenkamp

The Construct of Optimum Stimulation Level

The notion that human behavior is sometimes instigated by the mere desire to attain a satisfactory level of stimulation has figured prominently among psychological theories investigating motivational tendencies as causes of people's actions (Berlyne, 1963; Zuckerman, 1994). People tend to prefer intermediate levels of stimulation, referred to as the *optimal stimulation level* (*OSL*) in the literature. There are reliable individual differences in the amount of stimulation considered optimal by a given person, which appear to be affected by social learning factors as well as biochemical substances (Zuckerman, 1994).

To attain a satisfactory level of stimulation, a person may engage in exploration of the environment. As stated by Berlyne (1963, p. 288), exploratory behavior is "an end in itself" "with" the sole function of changing the stimulus field." Psychologists have studied exploratory tendencies extensively, and the general finding has been that people with higher OSLs engage in exploratory behaviors to a greater extent than people with lower OSLs (see Zuckerman (1979, 1994) for reviews).

Measurement. Several self-report measures have been developed to assess OSL, among them the 40-item Arousal Seeking Tendency (AST) scale, and its revised version, the 32-item AST-II scale, the 95-item Change Seeker Index (CSI), the 40-item Sensation Seeking Scale, version V (SSS-V), and the 80-item Novelty Experiencing Scale (NES). Despite the different labels, all scales load on the underlying construct of OSL (Steenkamp and Baumgartner, 1992). AST-I/II, SSS-V, and NES scales specify subfactors, but research has indicated that their factorial structure is generally unstable. Consumer researchers have, therefore, almost invariably used summated scores on the total scale. The reliability of the summated scores on each scale is high.

Steenkamp and Baumgartner (1992) concluded that CSI is the preferred instrument, based on extensive psychometric and nomological tests. However, with 95 items, CSI is much too long to be of practical use in most research applications. To address this issue, Steenkamp and Baumgartner (1995) developed a 7-item short-form version of the CSI scale. The 7-item scale has better psychometric properties and nomological validity than the original, 95-item scale. These items are listed in Table 1.

Relevance of Optimum Stimulation Level to Understanding Consumer Behavior

The consumer behavior literature has been dominated by the information processing paradigm, in which it is assumed that the consumer purposefully solves problems in order to achieve goals. However, since the seminal article by Holbrook and Hirschman (1982), researchers have increasingly recognized that many consumer behaviors are not (solely) purposeful in this sense, but also contain a strong exploratory component. For example, a consumer may not only buy a new product because she/he believes that the product is of better quality (information processing perspective) but also because she/he likes the excitement of trying out something new (exploratory perspective). As such, the exploratory perspective adds to our understanding of consumer behavior.

The fact that psychological research has uncovered that exploratory tendencies are related to a person's characteristic need for stimulation suggests that OSL may be a major determinant of consumer behaviors with strong exploratory elements. Indeed, a growing body of research has shown that OSL is an important factor in explaining a wide variety of consumer behaviors with an exploratory component. Consumers seeking thrills, adventure, disinhibition, new experiences, fantasies, cognitive or sensory stimulation, escape from boredom, and alternation among familiar things have been identified as engaging in exploratory consumer behaviors in order to raise their level of stimulation in life (Baumgartner and Steenkamp, 1996; Celsi, Rose, and Leigh, 1993; Holbrook and Hirschman, 1982; Raju, 1980; Steenkamp and Baumgartner, 1992; Steenkamp

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Items of CSI short-form scale.

1. I like to continue doing the same old things rather than trying new and different things.[a]
2. I like to experience novelty and change in my daily routine.
3. I like a job that offers change, variety, and travel, even if it involves some danger.
4. I am continually seeking new ideas and experiences.
5. I like continually changing activities.
6. When things get boring, I like to find some new and unfamiliar experience.
7. I prefer a routine way of life to an unpredictable one full of change.[a]

[a]This indicates reverse-coded item. Items are rated on a 5-point scale, ranging from −2 (completely false) to +2 (completely true) or from 1(completely disagree) to 5 (completely agree).

and Burgess, 2002; Steenkamp, ter Hofstede, and Wedel, 1999).

High OSLs have a greater preference for emotionally charged stimuli such as fear-arousing ads (Steenkamp, Baumgartner, and Van der Wulp, 1996) and have a greater interest in pursuing fantasies and fun (Holbrook and Hirschman, 1982). They engage more often in information search out of curiosity, generate more curiosity-based thoughts when exposed to ambiguous ads, and experience greater tedium during repeated exposure to the same ad (Baumgartner and Steenkamp, 1996; Steenkamp and Baumgartner, 1992). High OSLs exhibit more variety seeking and have a greater interest in knowing about novel or complex products and brands out of curiosity (Raju, 1980). OSL is related positively to a person's tendency to try out new retail outlets (Mittelstaedt *et al.*, 1976) and to his/her willingness to purchase new products and brands (Baumgartner and Steenkamp, 1996; Gielens and Steenkamp, 2007; Steenkamp and Gielens, 2003). Steenkamp and Burgess (2002) provide evidence that OSL does not only have relevance for understanding behavior of Western consumers but also holds great promise for understanding the behavior of consumers in emerging markets.

Conclusion

OSL is firmly grounded in psychological theory and can be measured reliably with self-report measures. OSL is of great relevance for understanding consumer behavior, both in high-income countries and in emerging markets. While it is generally believed that general personality traits have little explanatory value for consumer behavior, OSL is clearly an exception. Since many consumer behaviors have an exploratory component, the theoretical potential of OSL is substantial. Therefore, consumer researchers are encouraged to include OSL more often in their research design.

Bibliography

Baumgartner, H. and Steenkamp, J.-B.E.M. (1996) Exploratory consumer behavior: conceptualization and measurement. *International Journal of Research in Marketing*, **13** (2), 121–137.

Berlyne, D.E. (1960) *Conflict, Arousal, and Curiosity*, McGraw-Hill, New York.

Berlyne, D.E. (1963) Motivational problems raised by exploratory and epistemic behavior, in *Psychology: A Study of Science*, vol. **5** (ed K. Sigmund), McGraw-Hill, New York, pp. 284–364.

Celsi, R.L., Rose, R.L., and Leigh, T.W. (1993) An exploration of high-risk leisure consumption through skydiving. *Journal of Consumer Research*, **20**, 1–23.

Gielens, K. and Steenkamp, J.-B.E.M. (2007) Drivers of consumer acceptance of new packaged goods: an investigation across products and countries. *International Journal of Research in Marketing*, **24**, 97–111.

Holbrook, M.B. and Hirschman, E.C. (1982) The experiential aspects of consumption: consumer fantasies, feelings, and fun. *Journal of Consumer Research*, **9**, 132–140.

Mittelstaedt, R.A., Grossbart, S.L., Curtis, W.W., and Devere, S.P. (1976) Optimum stimulation level and the adoption decision process. *Journal of Consumer Research*, **3**, 84–94.

Raju, P.S. (1980) Optimum stimulation level: its relationship to personality, demographics, and exploratory behavior. *Journal of Consumer Research*, **7**, 272–282.

Steenkamp, J.-B.E.M. and Baumgartner, H. (1992) The role of optimum stimulation level in exploratory consumer behavior. *Journal of Consumer Research*, **19**, 434–448.

Steenkamp, J.-B.E.M. and Baumgartner, H. (1995) Development and cross-cultural validation of a short form csi as a measure of optimum stimulation level. *International Journal of Research in Marketing*, **12**, 97–104.

Steenkamp, J.-B.E.M., Baumgartner, H., and Wulp, E.V. (1996) Arousal potential, arousal, stimulus attractiveness, and the moderating role of need for stimulation. *International Journal of Research in Marketing*, **13**, 319–329.

Steenkamp, J.-B.E.M. and Burgess, S.M. (2002) Optimum stimulation level and exploratory consumer behavior in an emerging consumer market. *International Journal of Research in Marketing*, **19**, 131–150.

Steenkamp, J.-B.E.M. and Gielens, K. (2003) Consumer and market drivers of the trial rate of new consumer products. *Journal of Consumer Research*, **30**, 368–384.

Steenkamp, J.-B.E.M., ter Hofstede, F., and Wedel, M. (1999) A cross-national investigation into the individual and national cultural antecedents of consumer innovativeness. *Journal of Marketing*, **63**, 55–69.

Zuckerman, M. (1979) *Sensation Seeking: Beyond the Optimal Level of Arousal*, Lawrence Erlbaum, Hillsdale.

Zuckerman, M. (1994) *Behavioral Expressions and Biosocial Bases of Sensation Seeking*, Cambridge University Press, New York.

marketing and feminism

Elizabeth C. Hirschman

The development of feminist theory in marketing has an uneven history. During the 1990s researchers published some key papers and articles which helped gain a foothold for feminist thought (see e.g., Bristor and Fischer, 1993; Hirschman, l993; Stern, 1993).

This effort was expanded by a growing set of feminist researchers during the 1990s to incorporate ecofeminism (Dobscha and Ozanne, 2001), gender roles (Penalosa, 1994), gay and lesbian studies (Penalosa, 1996) and critiques of marketing ideology as essentially masculine (Penalosa, 1994). These diverse lines of research were consolidated in an edited volume (Catterall, Maclaran, and Stevens, 2000) which effectively summarized the state of theorization and application of feminism in marketing over the prior decade.

During the past decade, that is, 2001–2010, feminism has been expanded further into marketing areas such as environmentalism (Dobscha and Ozanne, 2001; Scott and Penalosa, 2006), queer studies (Scott and Penalosa, 2006), and critical theory (Maclaran *et al*., 2008). However, feminist research has not made the advances one would have expected, given its early theoretical potency and intrinsic applicability to a wide set of marketing phenomena. In particular, little attention has been directed toward the economic progress of women in marketing careers or, indeed, within the academy, itself (for an exception, see Penalosa, 2000). For example, women have served as editors of the *Journal of Consumer Research*. However, they have not acted in a similar capacity in *the Journal of Marketing*, *Journal of Marketing Research*, or *Journal of Public Policy and Marketing*.

A portion of this slowdown is likely attributable to the redirecting of research attention toward critical theory, generally, and even more recently toward transformative research agendas on the part of feminist marketing scholars (see e.g., Catterall, Maclaran, and Stevens (2005, 2006)). However, the larger part may be due to a "feminist generational gap."

The majority of productive feminist scholars in marketing matured as researchers during the early 1990s, a time period during which feminism played a more active role in academic discourse and public awareness. Over the past decade, while the ranks of women scholars in marketing have doubled, there has been little interest among these newcomers to apply feminist thought in their research agendas. Indeed, many may see little or no need for feminism on either a personal or academic level. Sadly, this viewpoint on their part is mistaken, as there are still ongoing gender discrepancies in salary, promotion, and the larger academic reward structure, as for example, journal editorships.

Further, as Scott (2009); Dolan and Scott (2009) has argued, there are many pressing issues for women in developing countries which tend to be given little attention by feminist researchers in Western Europe and North America. Among these are legal restrictions on women's rights to own property and work outside the home, very insufficient access to health care and education, and restrictions on travel. Recent initiatives to provide basic health care and employment opportunities have the potential to remake the lives of hundreds of millions of women around the world (Scott, 2009). Thus, there remains much important work to be done and a reawakening of interest in the relationship between feminism and marketing would be of great value to the field – and the lives of women – in many ways.

See also *sex in advertising; subcultures*

Bibliography

Bristor, J.M. and Fischer, E. (1993) Feminist thought: implications for consumer research. *Journal of Consumer Research*, **19**, 518–527.

Catterall, M., Maclaran, P. and Stevens, L.L. (1997) Marketing and feminism: a bibliography and suggestions for further research. *Marketing Intelligence and Planning*, **15** (7), 67–80.

Catterall, M., Maclaran, P. and Stevens, L.L. (eds) (2000) *Marketing and Feminism: Current Issues and Research*, Routledge Interpretive Marketing Research series, Routledge, London.

Catterall, M., Maclaran, P. and Stevens, L.L. (2005) Postmodern paralysis: the critical impasse on feminist perspectives on consumers. *Journal of Marketing Management*, **21** (5-6), 489–504.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Catterall, M., Maclaran, P. and Stevens, L.L. (2006) The transformative potential of feminist critique in consumer research. *Advances in Consumer Research, Association for Consumer Research*, **33**, 222–226.

Dobscha, S. and Ozanne, J.L. (2001) An ecofeminist analysis of environmentally sensitive women. *Journal of Public Policy and Marketing*, **20** (2), 201–214.

Dolan, C. and Scott, L.M. (2009) Lipstick evangelism: avon trading circles and gender empowerment in South Africa. *Gender and Development*, **6**, 203–218.

Hirschman, E. (l993) Ideology in consumer research, 1980 and 1990: a marxist and feminist critique. *Journal of Consumer Research*, **19** (4), 537–555.

Maclaran, P., Catterall, M., Stevens, L. and Hamilton, K. (2008) *Reinstating Wider Social Critique in Research on Gender and Consumer Behavior*, Association for Consumer Research conference on Gender Marketing and Consumer Behavior, Boston.

Penalosa, L. (1994) "Crossing boundaries, drawing lines": a look at gender trouble in marketing research. *International Journal of Research in Marketing*, **11**, 359–379.

Penalosa, L. (1996) We're here, we're queer and we're going shopping. *Journal of Homosexuality*, **31** (1/2), 9–41.

Penalosa, L. (2000) You've come a long way baby? Negotiating Feminism in the Marketing Academy in the US, in *Marketing and Feminism*, (eds M. Catterall, P. Maclaran and L. Stevens), Routledge, London, 39–50.

Scott, L.M. (2009) The double X economy. *Business at Oxford*, **10**, 13–17.

Scott, L.M. and Penalosa, L. (2006) Matriarchal marketing: a manifesto. *Journal of Strategic Marketing*, **15**, 309–323.

Stern, B.B. (1993) Feminist literary criticism and the deconstruction of ads. *Journal of Consumer Research*, **19** (4), 556–566.

consumer categorization

Barbara Loken

Categorization Constructs and Definitions

A *consumer category* is "a set of products, services, brands, or other marketing entities, states, or events that appear to the consumer, related in some way" (Loken, Barsalou, and Joiner, 2008). The types of categories studied in marketing include *taxonomic* categories such as *product* categories (e.g., "snack foods") and *brand* categories (e.g., "Coach"). Consumers also have *goal-derived* categories, which focus on a desired outcome such as "ways to chat with my friends", and might include telephone, Facebook, and email as category members (Ratneshwar *et al.*, 2001).

Category members have *graded structure;* they vary probabilistically in how representative or *prototypical* they are of the category. For example, Coach handbags are more typical and a better example of the Coach brand category than Coach sunglasses. Prototypical category members tend to be better liked, have more ideal or favorable category features, and are encountered more frequently, than atypical members (Loken and Ward, 1990; Veryzer and Hutchinson 1998; Viswanathan and Childers, 1999).

Category Stability and Flexibility

Research on consumer categories yields several conclusions (see Loken, Barsalou, and Joiner, 2008, for a review). First, categories are both *stable* and *flexible*. A consumer's concept of Crest as a prototypical toothpaste brand stays relatively stable over time, and consumers would generally agree that toothpaste (rather than "dental hygiene") is a "*basic level*" of categorization. Category membership and representation are also flexible, shifting with the context. For example, the prototypical beverages "at a football game" are different from those "at an opera intermission". Consumers' goals often motivate their categorizations (Ratneshwar *et al.*, 2001). The goal-relevant attribute of "sophistication" may drive category membership of beverages at the opera but not at football games. Category experts tend to show more category flexibility and greater use of subcategories than novices (Cowley and Mitchell, 2003), and consumers categorize differently depending on their cultural self-view as either Eastern or Western (Jain, Desai, and Mao, 2007).

Brand categories (e.g., Kashi products) are flexible in that atypical extensions (e.g., Kashi candy) "fit" the category better under some conditions than others. When consumers are in a good mood, or seeking variety or risk, they rate moderately atypical brand extensions as better category members. People accept more moderately atypical category members when a brand category is broad (e.g., Healthy Choice) than narrow (e.g., Campbell's). However, extremely atypical category extensions (e.g., Kashi electronics) are viewed as atypical regardless of the context. *Assimilation-contrast* processes also can explain flexibility of category members (cf. Schwarz and Bless, 2007).

Category Inferences

A second conclusion is that consumers use category information to make inferences about new category members. Consumers transfer beliefs and attitudes about a brand (e.g., the Coach brand is associated with "leather" and "designer status") to a new brand extension (e.g., a new Coach pink leather belt) when the brand extension is viewed as similar to, or a good fit for, the brand.

Third, consumers use information about new category members to make inferences and judgments about the overall category. A new BMW hybrid might increase consumers' beliefs that BMWs have "good gas mileage" or decrease beliefs that BMWs have "excellent performance". Failed extensions in near product categories and successful extensions in far product categories affect brands, particularly when other information is inaccessible.

Subcategorizations

Fourth, consumers use *subtypes* or subgroupings of categories to more finely differentiate a

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

category. For example, Apple *subbrands* might include MP3 players (iPod), handheld computers (iPhone), laptop computers (iMac), and musical software (iTunes). Marketers use subbrands to call attention to unique or innovative features, and these subbrands sometimes develop into strong categories of their own. When a subbrand has multiple members (e.g., multiple electric cars in the vehicle category), consumers are more likely to position a new product within this subbrand (Lajos *et al.*, 2009).

Fifth, consumers enjoy products more and *satiate* less quickly when they categorize at a lower or *subordinate* level (Redden, 2008). For example, enjoyment of jelly beans continues longer when the candy is categorized specifically (e.g., orange, cherry) than generally (e.g., jelly bean). Subcategorizations focus people's attention on differences between category members. Having multiple subcategories (e.g., on a store shelf) signals variety, which can increase consumers' sense of self-determination and overall satisfaction (Mogilner, Rudnick, and Eyengar, 2008) but can also bias choices (Fox, Ratner, and Lieb, 2005).

Acknowledgment

Thanks to Joe Redden for comments and suggestions on an earlier draft.

Bibliography

Barsalou, L.W. (1985) Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **11**, 629–654.

Cowley, E. and Mitchell, A.A. (2003) The moderating effect of product knowledge on the learning and organization of product information. *Journal of Consumer Research*, **30**, 443–454.

Fox, C.R., Ratner, R.K., and Lieb, D.S. (2005) How subjective grouping of options influences choice and allocation: diversification bias and the phenomenon of partition dependence. *Journal of Experimental Psychology: General*, **134** (4), 538–551.

Jain, S.P., Desai, K.K., and Mao, H. (2007) The influence of chronic and situational self-construal on categorization. *Journal of Consumer Research*, **34** (1), 66–76.

Lajos, J., Katona, Z., Chattopadhyay, A., and Sarvary, M. (2009) Category activation model: a spreading activation network model of subcategory positioning when categorization uncertainty is high. *Journal of Consumer Research*, **36** (1), 647–650.

Loken, B. (2006) Consumer psychology: categorization, inferences, affect, and persuasion. *Annual Review of Psychology*, **57**, 453–485.

Loken, B., Barsalou, L.W., and Joiner, C. (2008) Categorization theory and research in consumer psychology: category representation and category-based inference, in *Handbook of Consumer Psychology* (eds C.P. Haugtvedt, P.M. Herr, and F.R. Kardes), Psychology Press, New York.

Loken, B. and Ward, J. (1990) Alternative approaches to understanding the determinants of typicality. *Journal of Consumer Research*, **17**, 111–126.

Meyers-Levy, J. and Tybout, A. (1989) Schema congruity as a basis for product evaluation. *Journal of Consumer Research*, **16**, 39–54.

Mogilner, C., Rudnick, T., and Eyengar, S.S. (2008) The mere categorization effect: how the presence of categories increases choosers' perceptions of assortment variety and outcome satisfaction. *Journal of Consumer Research*, **35** (2), 202–215.

Ratneshwar, S., Barsalou, L.W., Pechmann, C., and Moore, M. (2001) Goal-derived categories: the role of personal and situational goals in category representations. *Journal of Consumer Psychology*, **10** (3), 147–158.

Redden, J.P. (2008) Reducing satiation: the role of categorization level. *Journal of Consumer Research*, **34** (5), 624–634.

Schwarz, N. and Bless, H. (2007) Mental construal processes: the inclusion/exclusion model, in *Assimilation and Contrast in Social Psychology* (eds D.A. Stapel and J. Suis), Psychology Press, Philadelphia, pp. 119–142.

Veryzer, R.W. and Hutchinson, J.W. (1998) The influence of unity and prototypicality on aesthetic responses to new product designs. *Journal of Consumer Research*, **24** (4), 374–393.

Viswanathan, M. and Childers, T.L. (1999) Understanding how product attributes influence product categorization: development and validation of fuzzy set-based measures of gradedness in product categories. *Journal of Marketing Research*, **36** (1), 75–94.

environmental consumer behavior

John A. McCarty, L. J. Shrum, and Tina M. Lowrey

Environmental consumer behavior, or more commonly, "green" consumer behavior, refers to a general class of behaviors (and their underlying processes) involved in purchasing, using, and disposing of products and services with the intention of improving the environment. A variety of activities can be considered under this umbrella. Probably the two most prominent ones are the extent to which consumers buy products that are environmentally friendly and the extent to which they dispose of product waste in an environmentally responsible manner. From a perspective of environmental activities, these two are related in that the disposal of waste (e.g., recycling of paper) provides some of the materials for the production of environmentally friendly products. Thus, in the materials life cycle (*see* FAMILY LIFE CYCLE) (Shrum, Lowrey, and McCarty, 1996), from the manufacturing of products and their packaging through consumer use to the disposal or recycling of the products and/or packaging, these two consumer actions (green buying and recycling) are the behaviors by which consumers can have the greatest potential impact on the environment.

Although a variety of polls over the last couple of decades show that consumers desire to become more "green" and are concerned about the environment, such findings are not always consistent with what consumers do (*see* ATTITUDE–BEHAVIOR CONSISTENCY), particularly in the area of buying green. For example, a 2007 global survey by McKinsey & Company (Bonini and Oppenheim, 2008) showed that 87% of the respondents indicated that they were concerned about the environment; however, only 33% reported that they bought or would buy green products. Thus, engaging in environmentally responsible behaviors, like other behaviors, is only partially determined by individuals' values, ATTITUDES, and intentions. There are a number of other factors that may facilitate or impede green behavior. In particular, polls typically show that consumers' reluctance to buy green relates to several key impediments: eco-friendly products are perceived as more expensive than other product offerings; eco-friendly products are perceived as less effective than other alternatives; and consumers are often skeptical of the claims made by environmentally friendly brands.

Besides buying green products, the other major way that consumers can act on their concern for the environment is to recycle the waste from the products they purchase. Like buying green, however, concern for the environment is but one of the factors that relate to the extent to which consumers will recycle materials such as paper, glass, and plastic. In terms of facilitating or inhibiting factors, the perceived convenience of recycling appears to facilitate the extent to which consumers recycle (McCarty and Shrum, 2001). Municipalities and organizations have caught on to this, and over the years there have been increased efforts to ease the burden on consumers to recycle. From the times when consumers had to take recyclables to special locations or bring recyclables to recycling drives at specific times, most cities now provide containers for recycling materials that are picked up at regular intervals.

Although making recycling more convenient increases recycling rates, it is still critical for consumers to believe that it is important. In fact, research has shown that consumers who feel that recycling is important tend to perceive it as less inconvenient, compared with those who do not feel it is important (McCarty and Shrum, 2001). The belief that recycling is important is driven, to some degree, by a variety of social and personality variables, such as values and attitudes; thus, these individual difference variables are critical to the success of consumer recycling.

Researchers have searched for antecedents of consumers' beliefs about the importance of recycling and pro-environmental behavior more generally. It appears that psychographic and personality variables show a stronger relationship with recycling attitudes than do demographic variables such as age and income (Straughan and Roberts, 1999). In particular, variables that are related to recycling include materialism (the extent to which one attaches importance to possessions), collectivism (the

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

extent to which one believes in the importance of the group over the individual), and locus of control (the extent to which people believe that they control their lives rather than that they are controlled by external forces). Those who are more materialistic tend to have more negative attitudes toward the environment than do those who are less materialistic (Kilbourne and Pickett, 2006). Those who are more collectivistic tend to believe that recycling is more important than do those who are less collectivistic (McCarty and Shrum, 2001). People with an internal locus of control tend to believe that recycling is more important than do those who have an external locus of control (McCarty and Shrum, 2001). Presumably, those with an internal locus believe that their environmental efforts will have an impact and thus believe that such activities can be important.

Bibliography

Bonini, S. and Oppenheim, J. (2008) Cultivating the green consumer. *Stanford Social Innovation Review*, **6**, 56–61.

Kilbourne, W. and Pickett, G. (2006) How materialism affects environmental beliefs, concern, and environmentally responsible behavior. *Journal of Business Research*, **61**, 885–893.

McCarty, J.A. and Shrum, L.J. (2001) The influence of individualism, collectivism, and locus of control on environmental beliefs and behavior. *Journal of Public Policy and Marketing*, **20**, 93–104.

Shrum, L.J., Lowrey, T.M., and McCarty, J.A. (1996) Using marketing and advertising principles to encourage pro-environmental behaviors, in *Marketing and Consumer Research in the Public Interest* (ed. R.P. Hill), Sage Publications, Thousand Oaks, pp. 197–216.

Straughan, R. and Roberts, J. (1999) Environmental segmentation alternatives: a look at green consumer behavior in the new millennium. *The Journal of Consumer Marketing*, **16**, 558–575.

habit in consumer behavior

Leona Tam

A habit is a behavioral disposition in which past responses are triggered directly by associated context cues. Consumer habits can be triggered by a wide array of cues in purchase and consumption contexts, including physical settings, presence of others, mood, time of day, and actions that immediately precede a habitual response. Thus, the theater environment may trigger popcorn consumption, or a visit from family around dinnertime might trigger a trip to Olive Garden.

How is habit formed? Consumer habits are formed when purchases and consumptions are repeated under similar circumstances. Such repetitions strengthen the cognitive associations between consumer responses and purchase and consumption contexts. When habits have formed, perception of the familiar cues activates a representation of the practiced, habitual response in memory. Given this reliance on cues, consumer habits are maintained when cues in the purchase and consumption contexts remain stable. Habits change when the cues are altered or when consumers are motivated and able to inhibit their automatic responses to the cues.

Even when consumers intend to try new brands and thereby change their established brand habits, they may continue to repeat past purchase and consumption activities when those responses are cued automatically by stable features of the environment. Because habits are automatically brought to mind in the appropriate circumstances, they tend to be repeated unless consumers are willing and able to exert effortful control to inhibit the activated response and to select a new one, or when the circumstances triggering habitual responses are changed.

Owing to the automaticity in consumer habits, consumers' habits do not necessarily change through altering behavioral intentions. How can consumers change their habits? One possibility is to take advantage of the *habit discontinuity effect*, or the disruption in habits that naturally occurs when people change life contexts. Change in performance contexts can shift consumer habits out of an automatic mode so that purchase and consumption become more deliberate and guided by consumer intentions. With changes in performance contexts, the habitual behavior may no longer be automatically brought to mind, and people may be released to act in ways consistent with their intentions and preferences. Marketers could target interventions to change consumers' habits to use competing brands for times at which consumers undergo naturally occurring changes in habit contexts, for example, when consumers relocate and may not have access to old retail outlets and associated brands. Retailers like Home Depot and Bed Bath & Beyond offer discounts as incentives to encourage patronage of consumers who have recently relocated. This strategy should effectively attract habitual consumers of competitor's brands whose habits are disrupted because of relocation.

Habitual consumers can also undertake to change their response to context cues, but not without sufficient regulatory strength to inhibit habits. In the implication of the willpower required for habit inhibition, marketers may wish to set up promotions and other sales events so that consumers repeatedly purchase and use their brand in ways that foster habit formation. Once formed, habits persist because they require willpower effort to inhibit. Illustrating such a promotion, a consumer recently received in the mail six $10 coupons for a new grocery store. Each coupon was valid for one week, and the coupon promotion was held over a period of six consecutive weeks. In order to use the coupons, she went to that grocery store every week, despite that there was another store she preferred. After six weeks, when all the coupons were used, she found herself going automatically to the new grocery store for the next shopping trip. She had formed a habit to frequent the new store, and, especially after a depleting day at work, it was easier to follow that habit than to inhibit it and choose another store.

Consumer habits are sometimes referred to as low involvement decision making, given the lack of effortful processing in consumer habit performance. However, consumers often repeat important purchases and consumptions and thus form habits. Health conscious consumers purposefully go to a salad bar for lunch on weekdays and are very likely going to form

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

a strong habit to do so after a few weeks of repetitions. Therefore, consumer habits are not limited to products of low involvement to consumers. The critical distinction between habits and brand loyalty for marketers lies in whether repeated purchase and consumption reflect the repetition of a particular response (habits) or whether they reflect an evaluative disposition that can guide repetition of a variety of brand responses (loyalty). Thus, for marketers, managing habitual customers requires different programs than those that are typically used to manage loyalty. Managing habits involves changing or maintaining the cue or the response to a cue, while managing loyalty involves improving or changing consumers' brand evaluation or favorability.

See also *consumer behavior analysis; consumer brand loyalty; implicit consumer cognition; self-regulation*

Bibliography

Beatty, S.E. and Kahle, L.R. (1988) Alternative hierarchies of the attitude-behavior relationship: the impact of brand commitment and habit. *Journal of the Academy of Marketing Science*, **16** (2), 1–10.

Ji, M.F. and Wood, W. (2007) Purchase and consumption habits: not necessarily what you intend. *Journal of Consumer Psychology*, **17** (4), 261–276.

Johnson, E.J., Bellman, S. and Lohse, G.L. (2003) Cognitive lock-in and the power law of practice. *Journal of Marketing*, **67** (2), 62–75.

Kaas, K.P. (1982) Consumer habit forming, information acquisition, and buying behavior. *Journal of Business Research*, **10** (1), 3–15.

Khare, A. and Inman, J.J. (2006) Habitual behavior in American eating patterns: the role of meal occasions-link. *Journal of Consumer Research*, **32** (4), 567–575.

Murray, K.B. and Häubl, G. (2007) Explaining cognitive lock-in: the role of skill-based habits of use in consumer choice. *Journal of Consumer Research*, **34**, 77–88.

Tam, L., Bagozzi, R.P. and Spanjol, J. When planning is not enough: the self-regulatory effect of implementation intentions on changing snacking habits. *Health Psychology*. (in press)

Tam, L., Wood, W. and Ji, M.F. (2009) Brand loyalty is not habitual, in *Handbook of Brand Relationships* (eds J. Priester, D. Maclnnis and C.W. Park), ME Sharpe.

Verplanken, B., Walker, I., Davis, A. and Jurasek, M. (2008) Context change and travel mode choice: combining the habit discontinuity and self-activation hypotheses. *Journal of Environmental Psychology*, **28** (2), 121–127.

Wirtz, J., Mattila, A.S. and Lwin, M.O. (2007) How effective are loyalty reward programs in driving share of wallet? *Journal of Service Research*, **9** (4), 327–334.

Wood, W. and Neal, D.T. (2007) A new look at habits and the habit-goal interface. *Psychological Review*, **114** (4), 843–863.

Wood, W., Quinn, J. and Kashy, D. (2002) Habits in everyday life: thought, emotion, and action. *Journal of Personality and Social Psychology*, **83**, 1281–1129.

Wood, W., Tam, L. and Guerrero Witt, M. (2005) Changing circumstances, disrupting habit. *Journal of Personality and Social Psychology*, **88** (6), 918–933.

childhood socialization and intergenerational influences

Elizabeth S. Moore

"To what extent do family influences experienced during childhood continue to exert influence on our behavior as adults?" This question frames the study of *intergenerational influence*, which is defined as the transmission of information, resources, and beliefs from one family generation to the next. Given the prolonged nature of childhood learning, the fact that some beliefs and attitudes formed within the household persist into adulthood comes as no surprise. It is the *nature* of these influences that is of interest, a topic rooted in socialization theory, and a long-standing area of inquiry in the social sciences (Macoby, 2007).

Socialization is the process through which individuals develop skills, beliefs, and behaviors needed to function successfully in society. It is a primary mechanism by which culture sustains itself. It also helps individuals to develop their personal identities, and assume new roles as they mature. Although socialization is a life-long process, childhood and adolescence are particularly crucial periods of this process.

The family, schools, mass media, religious institutions, and peers contribute in important ways to the socialization process. However, the family is the first and typically most powerful socialization agent, with parents and other family members serving as channels of information, sources of social pressure, and support. Socialization is a bidirectional process with parents and children influencing one another.

Consumer socialization, defined as "processes by which young people acquire the skills, knowledge and attitudes relevant to their functioning in the marketplace" is one aspect of the broader socialization process (Ward, 1974). Published research provides insight into the development of children's decision-making skills and marketing knowledge, particularly advertising (see John's, 1999 review). It also helps to understand parent–child interaction in household purchase decisions and impacts of parenting style (Carlson and Grossbart, 1988).

Intergenerational consumer research emerged in the 1970s. Since then scholars have shown that these effects are significant, interesting, and potentially important in the marketplace. Early research focused on consumer buying styles. Hill's (1970) longitudinal study showed, for example, that financial planning skills transfer across generations, particularly in families that are poor planners. Parent–child commonalities have also been discovered for choice heuristics, marketplace beliefs, and innovativeness. Interestingly, Cotte and Wood (2004) found that consumer innovativeness is affected by both parents and siblings, but parents are more influential.

Researchers have also examined the transmission of brand and product preferences, outcomes with direct strategic implications. For example, the public versus private nature of consumption for shopping goods may impact family influence levels (Childers and Rao, 1992). Intergenerational impacts are also a source of brand equity for packaged goods, but only for some brands, thus posing strategic challenges for particular marketers (Moore, Wilkie, and Lutz, 2002). These effects are multidimensional (impacting usage, consideration, and preference) and difficult to isolate analytically. Further, many factors are at work in adulthood to sustain or disrupt them.

Current knowledge of intergenerational influences in consumer research is grounded in a relatively small research base. Primary emphasis has been on documenting effects and the forms these take (Moore and Wilkie, 2005). Intriguing challenges lie ahead in the development of new measurement approaches, in understanding the determinants of intergenerational influences, and variation across families.

Bibliography

Carlson, L. and Grossbart, S. (1988) Parental style and consumer socialization of children. *Journal of Consumer Research*, **15**, 77–94.

Childers, T.L. and Rao A.R., (1992) The influence of familial and peer-based reference groups on consumer decisions. *Journal of Consumer Research*, **19**, 198–211.

Cotte, J. and Wood, S.L. (2004) Families and innovative consumer behavior: a triadic analysis of sibling and parental influence. *Journal of Consumer Research*, **31**, 78–86.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Hill, R. (1970) *Family Development in Three Generations*, Schenkman, Cambridge.

John, D.R. (1999) Consumer socialization of children: a retrospective look at twenty-five years of research. *Journal of Consumer Research*, **26**, 183–213.

Macoby, E.E. (2007) Historical overview of socialization research and theory, in *Handbook of Socialization* (eds J.E. Grusec and P.D. Hastings), Guilford Press, New York, pp. 13–41.

Moore, E.S. and Wilkie, W.L. (2005) We are who we were: intergenerational influences in consumer behavior, in *Inside Consumption: Frontiers of Research on Consumer Motives, Goals and Desires* (eds S. Ratneshwar and D.G. Mick), Routledge, London, pp. 208–232.

Moore, E.S., Wilkie, W.L., and Lutz, R.J. (2002) Passing the torch: intergenerational influences as a source of brand equity. *Journal of Marketing*, **66**, 17–37.

Ward, S. (1974) Consumer socialization. *Journal of Consumer Research*, **1**, 1–14.

consumer brand loyalty

Richard L. Oliver

Introduction

This article describes the general field of consumer (customer) loyalty as it is currently conceptualized. Whereas the field has been approached from many perspectives, the most common appearing in the area of repeat purchasing, more psychologically based approaches are now coming into fruition.

Here, the underlying mechanisms of how consumers manifest, consciously or subconsciously, their loyalty mind-set are explored so that greater loyalty development and switching avoidance can be engendered. Readers interested in more detail and elaboration, including discussion of topics not covered here, should consult the author's original works (Oliver, 1997; Oliver, 2010) and an earlier conceptual article on loyalty determinants (Oliver, 1999).

Consumer Brand Loyalty

Oliver (1997, p. 392), had proposed the following definition as being consistent with conceptual and empirical evidence:

Customer loyalty is a deeply held commitment to rebuy or repatronize a preferred product or service consistently in the future, *despite* situational influences and marketing efforts having the potential to cause switching behavior (italics in original).

To put this consumer display of loyalty in perspective, an elaboration of the stages and their vulnerabilities, which will address the "despite" clause, brings what is known of loyalty to the recognition of a social extension, discussed later in this article.

Loyalty phases. The framework used here follows the cognition–affect–conation historical pattern of attitude engagement (Taylor, Hunter and Longfellow, 2006), but differs in the sense that consumers can become "loyal" or "locked" at each loyalty phase. Specifically, consumers are thought to first become loyal in a cognitive sense, then later in an affective sense, still later in a conative manner, and finally in a behavioral manner, described as *action inertia*.

Cognitive loyalty. In this first loyalty phase, the brand attribute information available to the consumer suggests that one brand is preferable to its alternatives. This stage is referred to as *cognitive* loyalty or loyalty based on brand belief only. Cognition can be based on prior or vicarious knowledge and recent experience-based information. Loyalty at this phase is directed toward the brand because of this "information" (attribute performance levels). This consumer state, however, is of a very shallow nature. If the transaction is routine, so that satisfaction (*see* CUSTOMER SATISFACTION) is not processed (e.g., trash pickup, utility provision), the reaches of loyalty are no deeper than mere performance, and "performance-only" models apply here. *If* satisfaction is processed, it becomes part of the consumer's experience and begins to take on affective overtones.

Affective loyalty. At the second phase of loyalty development, the emergence of a liking or attitude toward the brand is required on the basis of cumulatively satisfying usage occasions. Commitment at this phase is referred to as *affective* loyalty (*see* EMOTION) and is encoded in the consumer's mind as cognition *and* affect. Whereas cognition is directly subject to counterargumentation, affect is not as easily dislodged. Here, the brand loyalty exhibited is directed at the degree of affect for the brand. Like cognitive loyalty, however, this form of loyalty remains subject to switching as evidenced by the data showing that large percentages of brand defectors claim to have been previously satisfied with their brand, a phenomenon known as the *satisfaction trap* (Reichheld and Teal, 1996). Thus, it would be desirable if consumers were loyal at a deeper level of commitment.

Conative loyalty. The next phase of loyalty development is the *conative* stage, as influenced by repeated episodes of positive affect toward the brand. *Conation*, by its definition, implies a brand-specific commitment to repurchase. Conative loyalty, then, is a loyalty state containing what at first appears to be the deeply held commitment to buy, noted in the loyalty definition. However, this commitment is to one's intention to rebuy (*see* CONSUMER INTENTIONS) the brand and is more akin to

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

motivation. In effect, the consumer desires to repurchase, but like any "good intention," this desire may be an anticipated but unrealized action. At this point, commitment is an additional overlay to the prior cognitive and affective bases of loyalty.

Action loyalty. The study of the mechanism by which intentions are converted to actions is referred to as *action control* (Kuhl and Beckmann, 1985). In the action control sequence, the motivated intention in the previous loyalty state is transformed into *readiness to act*. The action control paradigm proposes that this is accompanied by an additional *desire to overcome obstacles* that might prevent the action. Action is seen as a necessary result of engaging both of these states. If this engagement is repeated, an "action inertia" develops, thereby facilitating repurchase.

Readers will note the correspondence between the two action control constructs, readiness to act and the overcoming of obstacles, to the loyalty definition presented earlier. Readiness to act is analogous to the "deeply held commitment to rebuy or repatronize a preferred product/service consistently in the future," while "overcoming obstacles" is analogous to rebuying "despite situational influences and marketing efforts having the potential to cause switching behavior." Thus, completing the earlier cognitive–affective–conative frameworks with a fourth, or action phase, brings the attitude-based loyalty model to the "behavior" of interest – the action state of inertial rebuying (*see* ATTITUDE–BEHAVIOR CONSISTENCY).

Obstacles to Loyalty: Switching Incentives

It may have occurred to the reader that true loyalty is, in some sense, fragile as long as alternatives are available. Competitors could (and do) take advantage of this position, engaging consumers via persuasive messages and incentives with the purpose of attempting to lure them away from their preferred offering. These verbal and physical enticements are the obstacles that brand or service loyalists must overcome. As may be evident at this point, the easiest form of loyalty to break down is the cognitive variety; the most difficult is the action state. Thus, the cognitive-to-action loyalty sequence brings the analysis closer to the emergence of "full" loyalty, but still may fail as each phase is subject to attack, as follows.

Cognitive loyalty. In the case of cognitive loyalty, it has been noted that this level of loyalty is based on purely functional characteristics, primarily costs and benefits, and is thus subject to functional shortfalls. For example, in many areas, it has been shown that deteriorating performance, apart from dissatisfaction, is a strong enhancement to switch. Thus, cognitive loyalty is actually "phantom loyalty," as it is directed to benefits and costs and not to the brand. Costs (including prices), in particular, are a major component of brand switching at the cognitive level.

Before moving to affective loyalty, it is necessary to address the commonly observed breakdown in the satisfaction–loyalty link (i.e., the "trap"). Two situations are of note. The first is apparent disloyalty (switching) in the face of satisfaction and the second is apparent loyalty when encountering very low levels of satisfaction (dissatisfaction). The first is easily explained by cognitive loyalty that is overcome by attractive alternatives, as noted. The second is also easily explained by a phenomenon that has come to be known as *cognitive lock-in* (Büschken, 2004).

Lock-in can be achieved via many tactics. The most obvious is through a supply monopoly such as video game cartridges. At the interpersonal level, consumers have been known to become loyal to particular provider employees (e.g., salespeople). It is also known that normal consumer failings such as the sunk cost phenomenon and ordinary consumer learning are strong motivators to stay with an otherwise dissatisfying provider. Working in parallel with these influences is the notion of simple fear of the unknown. These are all reasons for firms to "ensnare" consumers so that they are held through "captive loyalty."

Affective loyalty. Barring lock-in, affective loyalty can become susceptible to dissatisfaction with the cognitive elements of a purchase, thereby inducing attitudinal shifts. A concurrent effect of dissatisfaction is the

increased attractiveness of alternatives, as noted. Thus, affective loyalty is first subject to the deterioration of its cognitive base, causing dissatisfaction, which has deleterious effects on the strength of one's attitude toward a brand, and hence, affective loyalty. It is also possible for competitive communications to use imagery and association to enhance the desirability of alternative brands while degrading that of the present brand.

Conative loyalty. Although conative loyalty brings the consumer to a stronger level of loyalty commitment, it has its vulnerabilities nonetheless. Although a consumer at this phase can weather some small number of dissatisfactory episodes, the consumer's motivation to remain committed can be "worn down" by barrages of competitive messages, particularly if they enhance the perceived severity of experienced dissatisfaction, a phenomenon called *prejudice*.

This occurs, not only via marketing communications, but also through social pressure from one's environment. In fact, it has been shown that social norms vastly dominate satisfaction (dissatisfaction) in the prediction of switching intentions (Roos, 1999). Additionally, competitive product trial, via samples, coupons, or point-of-purchase promotions, may be particularly effective here as the consumer has only committed to the brand, but has not committed to avoiding trial of new offerings. Thus, the conatively loyal consumer has not developed the "resolve" to intentionally avoid consideration of competitive brands.

At this juncture and perhaps before action loyalty manifests itself, the firm has achieved *product superiority*. Here, the firm has engendered enhanced liking – even an established preference – for its brand because of quality (information) and continued ability to satisfy. Additionally, the consumer is committed to its repurchase in the future. But, the consumer has not reached the state of resistance, resilience, and the overcoming of obstacles and adversity necessary for lasting loyalty to emerge.

Action loyalty and the beginning of fortitude. When reaching the action phase of brand attachment, however, the consumer has generated both the focused desire to rebuy a brand and only that brand, and has also acquired the skills necessary to overcome threats and obstacles to this quest. This consumer would be expected to routinely "tune out" competitive messages, to engage in effortful search for the favored brand, and to possibly even shun the trial of competitive brands. Marketers with action-loyal segments need not expend great sums on retention as, theoretically, their consumers would be governed by "inertial" repurchasing. Aside from deteriorating performance, a potential switching inducer at all stages, only insurmountable unavailability would cause this consumer to try another brand.

With the emergence of the action phase, it appears that the formula for loyalty has been largely crafted; later, this state will be referred to as *consumer fortitude*. It is intended that the action-loyal consumer has a deep commitment to repurchase, so much so that such behavior may be guiding itself in some habituated manner. But habituation is not the loyalty of interest. Rather, the quest is for a version of loyalty fitting the description of devotion "against all odds." When found, the state of consumer fortitude is achieved whereby the consumer fervently desires the consumable in a prohibitive, exclusive relationship. This should be a natural occurrence experienced by the consumer and not one created by the marketer.

It is to be borne in mind, however, that it is the province of competition to gain the consumer's attention so as to hear its communications. One major tactic by which this is accomplished, common in one or another way in all loyalty phases, is the creation of dissatisfaction with the current brand. In fact, the role of satisfaction in loyalty formation and defection can now be more fully specified. In the same way that satisfaction is a building block for loyalty, primarily at the affective loyalty stage, dissatisfaction is its Achilles tendon – for it is here that the competition can strike through the creation or facilitation of dissatisfaction at every stage.

Interpersonal loyalty: additional effects in services. Historically, discussion of loyalty effects has focused largely within the context of product marketing. With the possible exception of fan clubs, industries having a large service component were still thought to be governed by

loyalty to the core deliverable. Thus, Starbucks' loyalty is to its product and business model and not to its baristas (service providers). With recognition of the strong interpersonal component of services, loyalty now takes on additional dimensions of a much more binding and even overriding nature. Little research had been available in this new area, save for a small number of sources. This oversight has reversed itself with a number of insights beyond that of the more generic topic of service delivery, primarily with regard to service contact personnel (e.g., Price and Arnould, 1999).

Loyalty programs. We now discuss the emergence of *loyalty programs* that have sprung up since the time of early grocery store promotions. The issue now is that of program-based repetitive purchasing and its link to the mind-set of "true loyalty." One question that must be answered, then, is that of the role of mechanical incentives or rewards in loyalty apart from habit formation, but including inertial states that tend to persist until their exuberance wears off. It is not clear that loyalty programs are quite so "malleable."

The evidence for the success of such programs can be assessed along the four stages of progression. Do purchasing rewards pass the test of cognitive through action sequencing with the attendant behavioral building blocks? The answer given here is "no." Can reward programs be made to provide sufficient structure so that the consumer can become loyal to something such as loyalty to the program and loyalty to the behavior of pursuing its rewards? Possibly "yes." The interesting aspect of this analysis is that insights may be gained into the pursuit of "true" consumer loyalty.

Cognition in loyalty programs. Most rewards programs are based on lock-in principles. Consumers are given participation points that accumulate over time. Programs that give immediate rewards without long-term implications (such as discounts) do little other than condition the consumer to expect future discounts. Moreover, the reward must have value in and of itself or must add value to the basic consumable. Carnival prizes, for example, are too fleeting a reward to the consumer to have lasting value. In contrast, higher order values (e.g., attainment) do have lasting value either in terms of self-esteem or social approval. Thus, card levels (silver, gold, platinum) or club achievement identification (million dollar, million mile) may work. Tangibles (upgrades, free flights) work for the moment and are good future "carrots." But few of these have lasting implications for the firm.

Loyalty to the prize simply does not engender loyalty to the firm. The firm becomes an instrument, an enabler. Sports is a good example of the many forms of "loyalty to," and an industry that is now in the early throes of programmatic "loyalty." What is the object of the loyalty target? Is it the sport itself, the statistics (including win/loss), team identity, local spirit, the crowd (fans), becoming one of the fans, the camaraderie, or the new incentives that teams are deploying, including preferred seating, among others. And one should not forget the entertainment venue of music, one of the most addictive pursuits due to its biological origin.

Now, if sports and music need some level of minimal incentives, tangible or not, what of truly mundane consumables such as air travel? And what of ordinary everyday purchasing where credit card programs hold sway? The essential value of a reward is similar to all valued consumables. As promotions, they should take on the characteristic of a desired, wanted, or needed item. But simple wanting satisfies only at the cognitive level. Loyalty "icons" must give pleasure in anticipation, in receipt, in use, in ownership, and in further desire. If rewards are viewed as a collection or accumulation, this in and of itself must be of value.

Affect in loyalty programs. This discussion would be incomplete if the next level of affective loyalty were omitted. Does a program enhance the liking one expresses toward the firm? This is where loyalty programs begin to falter. Unless the program rewards can transfer the pleasure of the reward to the pleasure of consuming the sponsor's deliverable, the program becomes available to any competitor. And what if the program becomes the product? When consumers begin buying Crackerjacks for the prizes and discard the product, the product becomes the prize, of little essential value. People begin taking

flights for the next flight they will obtain or to complete the requirements for a free flight of even greater value or to stave off expiration.

The next two phases, conative and action loyalty, remain as goals, but are easily hidden or mimicked. In some sense then, the loyalty progression bifurcates at affective loyalty into a true psychological progression where the stages are in synchrony or into a false appearance of loyalty where consumers go through the motion of being committed and then repeat-purchase. At this point, without a psychological measurement metric, the mechanical and psychological loyalty patterns appear identical. From the firm's perspective, it won't be known which-is-which until true loyalty is allowed to emerge. This can be observed with either of any of a variety of "acid tests" such as termination of the program, comparison to a control market without the program, or lack of defections to competitors with equal or better programs.

The data are universally discouraging with a small number of exceptions (e.g., Dowling and Uncles, 1997). What is clear is that programs increase purchasing in the classic carrot and stick fashion. What is not clear is whether revenues net of the additional costs are profitable and these data are not available except anecdotally. It is also clear that heavy current purchasers are most favorably affected and benefit at the costs of less frequent purchasers. Besides, strategically, programs are necessitated, and costs incurred, when faced with close competitors with programs.

Consumption Communities

Three new perspectives on customer loyalty are proposed, stated as questions: (i) Can the consumer elect to be self-isolated from alternative consumable overtures so that competitive information is blocked or screened? (ii) Can the consumer be socially integrated in an exclusive environment that envelopes and directs the consumer's choices in a satisfying way? (iii) Can the consumer effect a self-identity that corresponds only to selected brand(s) *within* the environment? These issues speak to the "community" of loyalty (*see* BRAND COMMUNITY), singularly in the case of self-isolation, communally in the case of the environment, and both in the case of a preclusive lifestyle.

Dimensions of the framework. Picture a 2×2 table representing the dimensions on which these new issues are based. The vertical dimension reflects the degree of "individual fortitude," or the degree to which the consumer fights off competitive overtures based on his/her allegiance to the brand and not on marketer-generated "information." Despite an artificial break in this continuum into high (top row) and low (bottom row) categories, it is acknowledged that loyalty commitment develops with the advancement of stages in the four-phase model. At the lowest levels of fortitude, the consumer has only brand-related information. At the highest levels of fortitude, the consumer has developed the action inertia discussed previously, and has *also* developed a fierce defense against competitive encroachment that approaches "blind faith."

The horizontal dimension of the table would illustrate low and high phases of community and social support (McMillan, 1996). Here, the community provides the impetus to remain loyal either because it is enticing in a passive sense or because it proactively promotes loyalty. This dimension is crossed with that of individual fortitude so that the high-high cell contains the apex of loyalty and the low-low cell the weakest case of more vulnerable "loyalty," basic product superiority.

Product superiority, the weakest form of loyalty in this new framework, has already been discussed in cognitive, affective, conative, and action terms. This reflects the traditional view of loyalty as resulting from high quality and/or product superiority, both of which are believed to generate a strong sense of brand-directed preference. At some point in the cognitive–affective–conative–action chain, the consumer will cross the threshold from "low" consumer fortitude to "high," largely on the basis of the degree of immunization against competition that may have developed. The perspective taken here, however, provides further conceptual content in the high-fortitude (and low-social support) quadrant. In addition to the consumer's desire to rebuy on the basis

of superiority, this framework suggests that he or she will also wish to rebuy based on determination or "determined self-isolation," that is, the consumer desires an exclusive relation with the brand and does not wish to be "courted" by other suitors.

The low-fortitude, high- social-support quadrant represents "village envelopment," in that it is analogous to the ever-popular concept of "it takes a village." Here, the consumer is sheltered from outside influences, nurtured in the use of selected and protected brands, and provided integrated and routinely updated consumption systems. The common computer platform and networking environment supported by most businesses is an example of this concept. The distinguishing feature here is that the consumer is a passive acceptor of the brand environment.

Lastly, the high-high quadrant, referred to as *immersed self-identity*, contains the combined influences of fortitude and social support. Here, the consumer has intentionally targeted (or has been targeted by) the social environment because it is consistent with and supports the self-concept. In effect, the consumer "immerses" his or her self-identity in the social system of which the brand is a part. This is a synergistic situation and is self-sustaining. The consumer fervently desires the product or service association, affiliates with the social setting knowing that it will be supportive of this association, and, at the limiting extreme is rewarded by the social system for his or her patronage. Religious institutions are good exemplars of this situation, although other secular social settings are equally illustrative such as fan clubs and alumni organizations.

It should be noted that the defining characteristics of these new perspectives are not directly under the control of management, but can be facilitated by it. They go beyond the cognitive–affective–conative–action sequence, because they transcend it. They tap into the socioemotional side of loyal consumption and closely access its "meaning," as discussed next.

Self-isolation as a sustainer of loyalty. Crossing the threshold from a belief in product superiority to brand-directed determinism and personal fortitude is a somewhat nebulous process. The transitioning mechanism is not well understood, even for areas of life where determinism is frequently observed (e.g., romance, religion, politics). One potential threshold transition phase is that of attachment. Attachment is something more than satisfaction as it requires states of satisfaction reinforcements over time until the "glue is set." This is an apt description since glue, being a chemical/mechanical adhesive, has resilience until the breaking strength is breached. Brand attachments, particularly those of "human brands," are not necessarily monogamous, can withstand bouts of dissatisfaction, and display commitment without being "ultimate." Still, this consumer may very likely be immune from competitive overtures, is unlikely to be swayed from determined repurchasing, may defend the brand fiercely, and would most probably promote the brand to others.

The social organization: the village. In its pure form, the village is a social alliance whereby the primary motivation to become loyal on the part of each consumer is to be one with the group and the primary motivation of the group overseers is to please their constituency. In this situation, the consumer becomes a (willing) participant because of the attention provided by its members. In the limiting case, the product/service is not the "consumable." Rather, it is the camaraderie provided by the social organization. Good examples of this are the numerous varieties of clubs including internet communities of all types.

This concept goes by many names in its various literatures, but is perhaps best exemplified as the aforementioned *consumption community*, on the basis of the widely observed notion that individuals feel a sense of community when they share the same consumption behaviors. More to the point, when they espouse and own the same brand, these groupings are known as *brand communities* (Muniz and O'Guinn, 2001). Note that the previous "club" examples are somewhat weaker forms of the social collective envisioned here as they largely assume only that the mere knowledge of shared consumption is sufficient to generate a consumption community. Thus, it appears that the social dimension of the proposed framework, much like the fortitude dimension, is a continuum and some of the

examples given "drift" to either the weaker or stronger side.

Individual and social integration: fully bonded loyalty. The final, high-high quadrant in the table encompasses a blend of personal identity with the cultural milieu surrounding the consumable. This situation is distinguished from the previous example of the village because, in the present case, the cultural/social environment may assume a passive or stationary, although enticing, role. Here, the consumer is drawn to the consumable environment as opposed to the situation where the environment defines consumption for the consumer, although this does occur. The main distinguishing feature of this cell is that the consumer finds a "natural match" with both the consumable and its environment.

This is a particularly healthy situation for the firm as the product/service is inextricably embedded within some portion of the consumer's psyche as well as his/her lifestyle. The consumable is now part and parcel with one's self-identity and with his/her social identity, that is, the individual cannot conceive of him/herself as whole without it. At the extreme, the object is present intensionally and extensionally. Here, the consumer would say that the object is "part of me" and that it is an "extension of me." He or she lives it. Strong examples include sects with religious overtones and cults, although consumables in the more ordinary consumption domain are clearly candidates. Common examples include products, services, and even images supported by fans with various levels of group identification. Typically, the identity of the consumer is not known to the team, artist, and so on. The allure of the larger consumption icon is sufficient to "hold" the consumer to the loyalty state.

This concludes the discussion of loyalty influences beyond the cognitive-to-action framework. A consumer's willingness to rebuy or repatronize cannot reach ultimate extremes until she/he is willing to adore and commit unfailingly to a product/service. Beyond this, the necessary additional adhesion stems from the social bonding of a consumption community and the synergy between the two. In essence, the consumer wants to be loyal, the social organization wants him/her to be loyal and, as a result, the two may become symbiotic. The role of satisfaction in loyalty formation, which has not been discussed in this section, is addressed next.

What is the relation between satisfaction and loyalty?. In Oliver (1999), a number of plausible relations were suggested linking satisfaction and loyalty. The appropriateness of these in light of the evidence offered is now discussed. The first, suggesting that satisfaction and loyalty are two manifestations of the same concept, is easily dismissed. From the many avenues of discourse presented and the definition of loyalty, it should be clear that the two concepts are distinctly separate. Satisfaction is a fairly temporal postusage state for one-time consumption or a repeatedly experienced state for ongoing consumption that reflects how the product or service fulfilled its purpose. Thus, satisfaction is *delivered* to the consumer. Loyalty, in contrast, is an *attained* state of enduring preference to the point of determined defense.

A second suggests that satisfaction is an essential ingredient for the emergence of loyalty, a third only that it is necessary, and another that there is an ambiguous overlap across the two. There is merit to these perspectives as no perspective discussed here entertains loyalty development without early or concurrent satisfying episodes. While it may be that satisfaction is not a core element of loyalty, particularly after loyalty has set, it is difficult to entertain loyalty development without satisfaction at some level, however small. The endurance of loyalty is another matter, however.

This brings this discussion to a dynamic perspective, which suggests that satisfaction becomes "transformed" into loyalty, after which they share virtually no common characteristics. This is truly an extreme position for it suggests that loyalty can never return to mere satisfaction. Indeed, some have suggested that there is a threshold at which loyalty can "revert" to *dissatisfaction* in the face of repeatedly unsatisfactory purchase episodes. Sometimes referred to as *catastrophe models*, discontinuous jumps are posited within an ordinarily continuous relationship, which is what many assume.

The reason for the ambivalence as to which conception is most accurate is that, even in the perspective taken here, there remain variants of loyalty. In addition to the cognitive-to-action sequence, it is suggested that there are now different degrees of loyalty depending on how many of the synergistic factors presented here are involved. Immersion-based loyalty is supported by the convergence of product, personal, and social forces and the consumer displaying this state has logical, personal, and communal loyalty sustainers. At the same time, competition is easily thwarted by these same forces. Removing any or one of these lowers the consumer's resistance to competitive persuasion. Loyalty supported only by the social environment permits the consumer to look beyond its borders. Loyalty supported only by fortitude is susceptible to "relapse" such as self-doubt, second thoughts, competitive onslaught, and repetitively unpleasant dissatisfactory experiences. Furthermore, as discussed throughout this article, loyalty supported only by product information is subject to competitive counterinformation.

Thus, the dynamic model comes closest to the perspective taken here except that satisfaction does not transform into loyalty as much as it is a seed requiring nurturance. These are the analogies to personal determination and social support. Without these additional factors, satisfaction, much like the seed, stays dormant. The consumer remains satisfied, but does not progress beyond that state. Even a flash of sustenance – like the flash of delight – will not begin the transformation process. Once the seed sprouts, it will grow if the requisites are there. Only the mature version contains the strength of survival despite lapses in its base form, namely, dissatisfaction.

The translation of loyalty to profits. The study of satisfaction as a precursor of loyalty, and exploration of the concept of loyalty itself, brings the discussion to what many considered to be the ultimate goal – profit. Until recently, insufficient data were available to establish the "much talked about" relationship between satisfied and loyal consumers and a firm's profits. While many working in the satisfaction field had assumed that the relationships between satisfaction and loyalty and between loyalty and profits are inherently intuitive and self-evident, others provide arguments to the contrary.

The reader may remember when stock market valuation models contained only financial and accounting data. They still do. It may seem strange indeed that the customer is not represented beyond mere sales (revenue) in these calculations. True, backward trends are analyzed and forecasts are made based on "hard" data. However, we now have equally hard data on customer satisfaction across industries and countries over time (Johnson *et al.*, 2001). Moreover, it is now recognized that the same customer satisfaction data statistically and significantly add to the explanation of changes (variation) in stock valuations (Morgan and Rego, 2006). Then what of loyalty?

Perhaps the greatest effect of loyalty on profit is the direct influence of a steady stream of future customers. In a manner similar to discounted cash flows, guaranteed future customers allow firms to budget offensive and defensive competitive efforts, to time tactical market moves, and to weather severe attacks by competitors until an appropriate strategy can be mapped. Moreover, since loyal customers, in the pristine case, require little, if any, marketing attention, efforts can be redirected toward product improvement and service enhancement. In the limiting condition, if a firm had an optimal number of perfectly loyal customers who, by definition, could not be swayed away, the firm's marketing costs would be zero. These findings no doubt will be buttressed by new work in the area as these conclusions are explored on a continuing basis.

Bibliography

Büschken, J. (2004) *Higher Profits Through Customer Lock-In: A Roadmap*, TEXERE (Thomson), Mason.

Dowling, G.R. and Uncles, M. (1997) Do customer loyalty programs really work? *Sloan Management Review*, **38**, 71–82.

Johnson, M.D., Gustafsson, A., Andreassen, T.W. *et al.* (2001) The evolution and future of national customer satisfaction index models. *Journal of Economic Psychology*, **22**, 217–245.

Kuhl, J. and Beckmann, J. (1985) *Action Control: From Cognition to Behavior*, Springer-Verlag, Berlin.

McMillan, D.W. (1996) Sense of community. *Journal of Community Psychology*, **24**, 315–325.

Morgan, N.A. and Rego, L.L. (2006) The value of different customer satisfaction and loyalty metrics in predicting business performance. *Marketing Science*, **25**, 426–439.

Muniz, A.M. Jr and O'Guinn, T.C. (2001) Brand community. *Journal of Consumer Research*, **27**, 412–432.

Oliver, R.L. (1997) *Satisfaction: A Behavioral Perspective on the Consumer*, Irwin/McGrawHill, New York, M.E. Sharpe, Armonk, NY.

Oliver, R.L. (1999) Whence consumer loyalty. *Journal of Marketing*, **63** (Special Issue), 33–44.

Oliver, R.L. (2010) *Satisfaction: A Behavioral Perspective on the Consumer*, 2nd edn., M.E. Sharpe, Armonk, NY.

Price, L.L. and Arnould, E.J. (1999) Commercial friendships: service provider–client relationships in context. *Journal of Marketing*, **63**, 38–56.

Reichheld, F.F. and Teal, T. (1996) *The Loyalty Effect: The Hidden Force Behind Growth, Profits, and Lasting Value*, Harvard Business School Press, Boston.

Roos, I. (1999) Switching processes in customer relationships. *Journal of Service Research*, **2**, 68–85.

Taylor, S.A., Hunter, G.L., and Longfellow, T.A. (2006) Testing an expanded attitude model of goal-directed behavior in a loyalty context. *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, **19**, 18–39.

hedonism: gratification is value

Barry J. Babin and Mitch Griffin

Mill's *Utilitarianism* impresses upon us the fact that gratification is tied to consequences. In other words, one finds value in activities because those activities may bring future contentment. Does anyone really want a root canal? Consumers seek this service because the end result will be greater contentment because of a lack of pain. While this notion is the received view of most consumer behavior – consumers are solving problems through their activities – the fact is that this utilitarian value in no way precludes the more nascent idea that value is also realized from the immediate gratification of these activities in and of themselves (Babin, Darden, and Griffin, 1994). An automobile is purchased because the consumer is able to address the task of transporting oneself, the family, and effects from one place to another. This utilitarian gratification will be realized in the future. However, the actual act of purchasing an automobile can also be directly gratifying. This is particularly true for the car enthusiast, but even a novice will likely feel some excitement taking a new (preferably expensive) car out for a test spin. This immediate happiness produces gratification, which is hedonic value.

The contrast between consumer activities characterized by hedonism, or experiencing things for the value of the experience itself, with activities performed because they are means to future ends, or workings of utilitarian value, provides a great deal of insight into consumers' behavior. For instance, consider the interplay between time and consumption. Although a topic that is not particularly well studied, most marketing management practice addressing time issues is directed toward reducing the time that it takes for a consumer to receive a service or accomplish a task. This is the utilitarian view. Hedonic value, in contrast, releases the consumer from the bounds of time in the sense that hedonism has no time cost. This does not mean that time does not pass, but the time invested now yields value because of the immediate gratification provided, rather than lacking or even diminishing consumer value.

One might ask, which of the two is greater, hedonism or utilitarianism? This is an empirical question (Babin and Attaway, 2000), but it also is a question that misleads to the extent that it casts these as opposing values. Marketers must provide experiences that provide high value and examples can be found where the provision is accomplished through predominantly utilitarianism (i.e., Wal-Mart) or hedonism (i.e., Disney). But perhaps the best route to truly great marketing is the provision of high utilitarian *and* high hedonic value. Consider how the modern day spa provides services a consumer might have previously received by visiting a hair salon, a nail salon, a massage therapist, a dermatologist, and so forth to receive the same amount of utilitarian value. But, these spas also provide an instantaneously gratifying experience. The consumer can experience hedonism. Fortunately for marketers, this added value is also associated with added payments by the consumer. When the two become inseparable, the experience of the extrinsic value provided by products and services provides intrinsic value in itself, the value experience is maximized.

Eventually, Mill (1861) argues that value is captured in contentment or happiness. Happiness eventually wins out as a greater and more interesting value being ephemeral yet powerful. People will sacrifice current contentment for greater happiness. Interestingly, marketing practice and theory devotes a great deal of attention toward satisfaction and less toward hedonism. Contentment is not nearly as motivating. The state of contentment itself is indicative of homeostasis and thus lacks motive power. However, hedonism leads to attachment and the desire to reexperience the gratification. Thus, perhaps consumer researchers would be well served to consider the relative power of value, particularly hedonic value, over satisfaction.

See also *brand value; consumer desire; customer lifetime value (CLV); customer solutions; emotion; international retailing; luxury brands versus other brand categories; variety-seeking*

Bibliography

Babin, B.J. and Attaway, J.P. (2000) Atmospheric affect as a tool for creating value and gaining share of customer. *Journal of Business Research*, **49**, 91–99.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Babin, B.J., Darden, W.R. and Griffin, M. (1994) Work and/or fun? Measuring hedonic and utilitarian shopping value. *Journal of Consumer Research*, **19**, 644–656.

Mill, J.S. (1861) in *Utilitarianism*, (ed. R. Crisp (1998)), Oxford University Press, Oxford.

variety-seeking

Aimee Drolet and Daniel He

Most of the choices consumers face are choices they face regularly and can be viewed fundamentally as being either consistent or inconsistent with their usual habits. On most of these occasions, consumers strive for consistency in their choices. Inconsistency can signal fickleness and fussiness or, even worse, moral and mental weakness. In contrast, consumers who display consistency in their choices are viewed positively, believed to possess personal strength and intellectual competency. Furthermore, consistency makes the task of choosing easier, as it takes less effort to make a choice that has been made before than to reconsider and choose differently. These are some of the benefits of behavioral consistency.

At the same time, the costs of inconsistency appear equally clear. Making the same choice repeatedly means experiencing the same result repeatedly. In one word, consistency is boring. Inconsistency in the form of choosing a variety of options can be stimulating and forestall satiation. It can also be perceived as creative and interesting by others. In short, sometimes inconsistency can be good, and consumers often embrace inconsistency by seeking to vary their behavior. Variety-seeking is a common consumer choice strategy.

The tendency to *variety-seek*, defined here as switching away from a choice made on a previous occasion, has been found in both interpersonal and intrapersonal consumer choice contexts. In interpersonal contexts, such as a group at a restaurant or bar, consumers will choose an option not chosen by other consumers in the group. They will order a different dish or drink even if their variety-seeking behavior will make them unhappier with their order (Ariely and Levav, 2000). And, in intrapersonal contexts, consumers will switch from choosing more-preferred options to less-preferred options despite lowered satisfaction (Simonson, 1990). Indeed, the tendency among consumers to seek variety in intrapersonal consumer contexts is pervasive enough that it extends beyond the choice of product options to the choice of product choice rules themselves (Drolet, 2002).

Why are consumers so motivated to seek variety in their choice making? Further, what factors moderate the tendency to seek variety? Specifically, what are the situational and person-specific factors that influence the degree to which consumers will endeavor to be consistent versus inconsistent in their choice behavior? This article addresses these questions.

Reasons Consumers Seek Variety

Researchers have offered several reasons why consumers are motivated to seek variety. Consumers seek variety for biological, utilitarian, and psychosocial reasons.

Biological reasons. At the most basic level, the tendency to variety-seek may be "hard-wired." Animal studies suggest that the desire to vary behavior is innate. Even after there is no new information to acquire through behavioral diversification, animals change their behavior, seemingly just for the fun of it. Likewise, people change their behavior because they can get bored with the repeating same experience over and over.

Smart producers of products respond to consumers' eventual boredom. For example, classical music, with its frequent repetitions and long movements, can sound tedious even to its own enthusiasts. The composer Joseph Haydn anticipated his own eighteenth century audience's boredom. He interspersed the gentle, melodic parts of the second movement in his Symphony No. 94 with sudden, thunderous bursts of music. Haydn later nicknamed the symphony the "Surprise Symphony" because the piercing notes he added grab listeners' attention with something different. They also wake up sleeping listeners.

Stimulation. Whether the products are songs (Ratner, Kahn, and Kahneman, 1999) or snacks (Simonson, 1990), consumers crave something different and choose options that allow them to experience something different. Change is inherently stimulating, and the inherent need for stimulation can be met by changing behavior. By choosing options that have not been chosen before or recently, consumers can increase their level of stimulation. Accordingly, by satisfying their need for

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

stimulation by variety-seeking in one product category, consumers' tendency to seek variety in another category is less (Menon and Kahn, 1995). Likewise, the need for stimulation can be met by having consumers make choices in complex environments (Menon and Kahn, 2002). In these environments, consumers are less likely to variety-seek.

Satiation. Besides boredom, satiation is another negative psychological consequence of repeated choice. Consumers become satiated with products, and variety-seeking among products is a way to stave off satiation (McAlister and Pessemier, 1982). Consumers also become satiated with product attributes. For example, in an experiment by Mitchell, Kahn, and Knasko (1995), consumers choosing among chocolate candies chose more variety in the presence of a stimuli-congruent scent (chocolate) versus a stimuli-incongruent scent (flowers).

Utilitarian reasons.

Uncertainty reduction. By choosing a mix of products with a mix of features, consumers can meet a mix of needs. Some needs may be known at the time of choice. However, some are not. Variety-seeking is a choice strategy that shields consumers from several sources of uncertainty. The future value of products and product features is one source of uncertainty consumers confront. For example, a consumer might not be able to predict the usefulness of a newly purchased product (e.g., GPS) or the usefulness of one of its features (e.g., voice activation). Another source of uncertainty is future preferences (Simonson, 1990). As consumers, we often do not know what our tastes will be in a year or even tomorrow. So, we cover our bets by choosing variety. We cast a wide net and widen our selection. By doing so, we are better able to cope with this source of uncertainty. Put differently, a strategy of variety-seeking has normative value if one assumes that, due to uncertainty about products and preferences, consumers are testing their utilities for choice options (McAlister and Pessemier, 1982) and choice rules over time (Drolet, 2002).

Simplify choice. Variety-seeking is one way that consumers cope with tough choices. Variety-seeking makes good sense for both cognitive and emotional reasons. In complex environments, for example, where there are large numbers of product options each described along large numbers of product features, consumers may not have the sufficient cognitive wherewithal to solve choice problems (Mitchell, Kahn, and Knasko, 1995). Variety-seeking is a choice heuristic that frees up cognitive resources. For example, consumers can focus on the relative value of attributes, an easier task than calculating attribute value trade-offs.

Choice problems can be emotionally difficult as well. Choosing one option means forgoing others. Consumers anticipate regret and seek to reduce it. Anticipated regret also drives consumers' unwillingness to make attribute value trade-offs that are emotionally threatening. To consumers, the idea of trading off attributes relevant to their self goals (e.g., personal safety and financial well-being) is an inherently aversive one. For consumers, variety-seeking is one way to avoid having to do so.

Maximize global utility. Variety-seeking makes normative sense if consumers are attempting to maximize global utility as opposed to local utilities (Kahn, Ratner, and Kahneman, 1997). When making a series of choices over time, consumers can choose to maximize the utilities associated with each choice made in the series (i.e., local utilities). Or, consumers can choose to maximize the utility associated with the entire series of choices (i.e., global utility). Maximizing local utilities presupposes that earlier choices have no effect on the satisfaction consumers experience from later choices. This may not always be the case, for example, because of satiation. However, maximizing global utility, that is the utility associated with the entire series of choices, may cause consumers to experience less overall satisfaction (Simonson, 1990).

Maximize retrospective utility. Variety-seeking also makes sense if consumers are attempting to maximize retrospective utility. Ratner, Kahn, and Kahneman (1999) provided evidence that this is sometimes the case. In their research, participants rated their memory for how much they enjoyed sequences of music that were more versus less varied. Participants' retrospective reports of their enjoyment were

higher for the more-varied (vs less-varied) sequence, even though their online reports of their enjoyment were lower for the more-varied (vs less-varied) sequence.

Psychosocial reasons.

Norm conformity. Some variation in choice behavior is viewed more favorably by others compared to no variation in choice behavior. A moderate amount of variation suggests positive values, such as discerning taste, and adaptive personality traits, such as flexibility. In contrast, the complete absence of variation suggests negative values, such as indiscriminant judgment, and maladaptive personality traits such as rigidity. When in the presence of others, consumers who follow a norm of variety-seeking are viewed more positively by others compared to those consumers who do not (Ratner and Kahn, 2002).

Need to appear unique. Another psychosocial reason that consumers seek to vary their choice behavior stems from the desire to be unique. In interpersonal contexts, a consumer will make different choices over others because these choices advance the goal of portraying oneself as unique. The consumer who makes the same choices as others runs the risk of appearing to imitate them. A study by Ariely and Levav (2000) supports this view. In their study, microbrewery patrons were offered the choice of one of the four free beer samples. In a "collective" condition, patrons ordered their samples one by one aloud to the waiter. Thus, each person in the group knew what each of the others ordered. In an "independent" condition, patrons wrote down their choice on individual menus. Thus, other persons in the group did not know what each person had ordered. Consistent with the view that consumers seek variety in order to portray themselves as unique, in both their own eyes and the eyes of others, patrons in the collective condition showed greater variety in their beer choices, suggesting they desired to be unique. They also expressed greater regret when asked if they wished they had made a different choice, suggesting that they had sacrificed their initial choice so as to be unique.

The variety-seeking in this study involved a sequence of individual choices. An implication of its results is that a consumer who is the first in a group to order can satisfy his or her own particular taste without having to worry about being perceived as unique since no one else in the group has chosen before him or her. Indeed, in follow-up tests, Ariely and Levav (2000) confirmed that consumers in the collective condition who were the first to order were just as satisfied as the consumers in the independent condition. Equal satisfaction implies that these consumers ordered the drink they most preferred and at the same time were able to appear unique. Consumers in the collective condition who were not the first to order may have appeared unique but were less satisfied with their drink order because their choice was contingent on the choices of others.

The desire to appear unique also plays a role in intrapersonal choice contexts. Behavioral change in oneself is associated with perceptions of oneself as unique. People who prize uniqueness separate themselves from others by maintaining a sense of "specialness." People can demonstrate specialness by choosing not to conform to standards of behavior set by others and also to standards of behavior they themselves set. Not conforming to one's own standards reveals a person to be unique as it represents a display of situational abnormality, one aimed at avoiding excessive similarity.

Moderators of Variety-Seeking Behavior

There are several situational and person-specific factors that moderate the tendency to seek variety and that determine the conditions under which this tendency is more versus less pronounced.

Situational factors.

Mood effects. Past research has shown that positive mood increases variety-seeking. For example, Kahn and Isen (1993) manipulated participants' mood to be positive by giving them a bag of candy. These participants chose more variety compared to participants in a control condition.

There have been multiple explanations offered for the increasing effect of positive mood on the tendency to variety-seek. One explanation is that positive mood is associated with sensation-seeking and novelty-seeking,

and variety-seeking is one way in which people can achieve these ends (Menon and Kahn, 1995). Another explanation is that positive mood increases heuristic processing, and variety-seeking is a choice heuristic that simplifies choice-making (Mitchell, Kahn, and Knasko, 1995). Negative mood is associated with systematic thinking and also monotonous behavior.

Relatedly, heuristic processing is associated with broadened thinking, whereas negative mood is associated with narrowed thinking. Kahn and Isen (1993) suggested that positive mood caused participants in their study to have a more optimistic view of consumption. Besides, positive mood's increasing effect on broadened thinking may lead participants to have a more flexible view as to what products fit in a category, thereby increasing the size of their consideration sets (i.e., the number of product options they would consider choosing).

Presence of others. The presence of a group forces individuals to deal with different goals, goals that can be achieved solely by the individual and goals that are accomplished by both the individual and the group (*see* SOCIAL INFLUENCE). Because these two goals often oppose one another, it may not be possible to meet both if each requires the choice of a different option. For example, when we are at a restaurant with friends, we must choose between these goals: we can either satisfy our own taste by ordering a meal without considering what our friends have ordered or we can portray ourselves as unique by ordering a meal contingent on the orders of others. If uniqueness is the goal, the recommended strategy would be to order an option that our friends do not. If uniqueness is not the goal, the recommended strategy would be to order an option that our friends order.

Ratner and Kahn (2002) demonstrated that the presence of others influences which kind of strategy consumers tend to pursue. Specifically, they found that consumers were more likely to variety-seek when their consumption was public versus private. Their finding implies that variety-seeking is a salient social norm and that compliance with this norm increases social evaluations. When their behavior is being observed, consumers will choose more variety than they would otherwise. By choosing a variety of products, consumers can appear more interesting, for example, a trait that is socially valued.

Physical confinement. Recent research by Levav and Zhu (forthcoming) found that a physically restrictive space causes consumers to seek more variety. A physically restrictive space arouses feelings of confinement and threatens personal freedom. As a result, consumers develop psychological reactance and try to reclaim their personal freedom by seeking variety. In their view, variety-seeking is a manifestation of the freedom to choose.

In one of their studies, participants were instructed to choose three out of ten candy bars from a bowl placed on top of a table. To get to the bowl, participants had to walk down either a wide aisle or a narrow aisle. They found that participants in the narrow aisle condition chose a greater variety of candy bars compared to participants in the wide aisle condition. In another experiment, participants were confined to either a narrow aisle room or a wide aisle room. They were asked to take a seat at the end of the room. Participants then learned about six charitable organizations – three more well-known and three less well-known – who were seeking donations. Participants were asked about their likelihood of donating to each. Participants in the narrow (vs. wide) aisle room were more likely to donate to the lesser-known charities. Some participants were also given the opportunity to make actual donations to the charities. Among those participants, the actual donation amount was higher in the narrow (vs wide) aisle condition. Levav and Rhu suggest that increased variety-seeking for the less well-known and more unique brands is the result of feelings of confinement and subsequent reactance to these feelings. Consistent with this suggestion, Levav and Rhu found that simply varying perceptions of confinement instead of actual physical confinement induces reactance and increases variety-seeking behavior.

Choice set size. An increase in the number of choice options can not only decrease a consumer's willingness to buy any option but can also lower the satisfaction with choices made from a larger set of choices. When the size of the

choice set increases, it is increasingly difficult for consumers to make a choice among items. One low-effort way for consumers to resolve difficult choice problems is to choose a mix of items. As mentioned, by choosing a mix of items, consumers can raise the odds of choosing a most preferred item and an item that can meet unknown future consumer needs.

Decision timing. The timing of decisions is an important situational determinant of the tendency to variety-seek. Consumers tend to seek more variety when making multiple product purchases simultaneously versus sequentially. An experiment by Simonson (1990) illustrates the effect of decision timing on variety-seeking behavior. Student participants were asked to choose from a set of snacks at the beginning of a class meeting. There were three conditions. In a sequential choice condition, participants made one choice each week for three weeks. In a simultaneous choice for sequential consumption condition, participants made all three choices at one time but received only one item each week. In a third condition, known as the simultaneous choices for immediate consumption condition, participants made three choices simultaneously and received all three snacks at the end of class that same day. Those in the simultaneous choices for sequential consumption condition were significantly more likely to choose a variety of snacks compared to those in the sequential consumption condition. Moreover, participants in the third condition who were given all three snacks the same day were even more likely to choose variety than participants in the other two conditions.

According to Simonson (1990), having to make multiple product purchases simultaneously versus sequentially forces the consumer to confront multiple sources of uncertainty, including the uncertainty due to the inability to predict future preferences accurately since preferences change over time as well as the uncertainty due to the inability to know which among a set of options is the most-preferred. With respect to the latter source of uncertainty, sequential choice making provides the consumer with the opportunity to update their preferences as each choice occasion passes.

Person-specific drivers.

Personality traits. Past research indicates that there are several personality traits associated with the tendency to seek variety versus avoid it. For example, the dispositional drive for stimulation predicts variety-seeking behavior. As mentioned above, a basic human need is the need for stimulation, and this need can be satisfied through varied behavior or the experience of novel stimuli. The extent to which stimulation is needed varies across individuals (*see* OPTIMUM STIMULATION LEVEL). Some people have a higher need for stimulation than others and are more likely to seek out new things. They do so by seeking variety.

Another personality trait that promotes the tendency of consumers to seek variety is the dispositional need for uniqueness. Consumers use rules to solve choice problems. Which rule consumers select from their repertoire of rules depends on various characteristics of choice problems. Past research has demonstrated that the rule consumers select is also independent of choice problem characteristics (Drolet, 2002). In particular, consumers who score high on the need for uniqueness scale tend to vary their selection of choice rules more than consumers who score low on the scale. Consumers who desire to see themselves as different from others can do so by avoiding repeated use of the same choice rule.

People vary in the dispositional tendency to be indecisive and this tendency is associated with increased variety-seeking behavior (Jeong, H.G., and Drolet, A. (2009) The indecisive consumer: delaying decision-making by seeking variety, unpublished research). Highly dispositionally indecisive individuals tend to view the prospect of choice as aversive. Choice-making for them is an experience filled with anxiety, worry, disappointment, and regret. Accordingly, they are reluctant to make choices and avoid doing so. As a result, choice is an activity that often ends in impasse and is deferred until later. In the meantime, the search for new alternatives and new information about alternatives can continue. Variety-seeking appears to be a strategy consumers use to delay a final choice and, as such, is a disproportionately attractive strategy for highly dispositionally indecisive consumers.

Given that chronic indecisiveness is associated with negative emotion, these results

appear at odds with research that has found an increasing effect of positive mood on variety-seeking behavior (e.g., Kahn and Isen, 1993). A study by Jeong, H.G., and Drolet, A. (2009) (The indecisive consumer: delaying decision-making by seeking variety, unpublished research.) reveals that negative mood increases variety-seeking among consumers who are high in dispositional indecisiveness but not among consumers who are low in dispositional indecisiveness. Chronically indecisive consumers report feeling better after choosing a variety of products.

Culture. In individualist cultures such as Western cultures, uniqueness has a positive meaning. In contrast, in collectivist cultures such as East Asian cultures, uniqueness has a negative meaning. Not following group norms of behavior undermines the cohesiveness of a group (*see* CROSS-CULTURAL PSYCHOLOGY OF CONSUMER BEHAVIOR). When maintaining the strength of a social group is important, and it is especially so for members of collectivist cultures, people must sometimes suppress their personal preferences in the interest of group harmony. In individualist cultures, where personal preferences often take precedence over social preferences, the act of choice is viewed as an act of self-expression. In collectivist cultures, the act of choice is not always associated with the self. These attitudes toward uniqueness and choice have important implications on the extent to which consumers seek variety. The tendency to variety-seek depends on consumers' cultural backgrounds.

Research by Kim and Drolet (2003) shows that consumers from individualist cultures tend to seek more variety compared to consumers from collectivist cultures. Specifically, in one study, US-born participants were more likely than Korean-born participants to vary the choice rules they used. Rule change behavior was not as apparent among Korean-born participants. Kim and Drolet hypothesized that this pattern of results is due to cultural differences in the value of uniqueness and the meaning of choice. In a different study, they used advertisements to prime uniqueness (or not) in both US-born and Korean-born participants. Consistent with their hypothesis, individuals exposed to advertisements with a uniqueness theme sought to vary their choice rule use more, regardless of cultural background. These findings raise the question of just how "hard-wired" the variety-seeking tendency is given that people's tendency to seek variety depends on cultural assumptions of choice and uniqueness.

Aging. There is some evidence that aging influences the tendency to seek variety depending on decision timing. In the research by Novak and Mather (2007), older adults chose more variety when making choices for immediate consumption than when making choices for future consumption. The amount of variety chosen by young adults did not differ when making choices for immediate versus future consumption.

Novak and Mather (2007) attribute this pattern of findings to older adults' increased emotion regulation abilities. As we age, we tend to focus more on positive information and less on negative information, and this change in focus influences choice. As a result, older adults tend to experience more positive mood. As discussed earlier, positive mood has been linked to increased variety-seeking behavior. If an initial experience is positive, and we would expect it to be more so for older adults who are better at focusing on positive aspects of experiences, an individual should be more willing to try out new things.

One reason older adults experience more positive mood is that they avoid potentially negative experiences. When not making choices in real time and there is no information as to whether the initial experience will be positive or not, older adults choose less variety. Again, there is uncertainty surrounding future preferences and this uncertainty is experienced as a negative one (Simonson, 1990). In the presence of this uncertainty, older adults might be more risk averse, seeking to avoid negative outcomes. They will gravitate toward the familiar. Consequently, they will seek less variety when making choices for future consumption.

CONCLUSION

Variety-seeking is a common consumer choice strategy. This article reviewed several reasons

why consumers view variety-seeking as a useful choice strategy. In addition, this article reviewed several moderators of the tendency to choose a variety of options. Both situational and person-specific factors influence the amount of variety consumers will choose.

In predicting the amount of variety consumers will choose, it is important to consider the interaction among such factors. For example, a person-specific driver like the personality trait need for uniqueness predisposes consumers to seek variety. However, this relationship depends on another person-specific variable, culture, which influences the degree to which a person values uniqueness. In summary, in order to predict the amount of variety consumers will choose, it is important to understand why some consumers are attracted to variety-seeking as a choice strategy whereas others are not. To do so, one must take into account the conditions under which some consumers are more or less likely to variety-seek than others.

Bibliography

Ariely, D. and Levav, J. (2000) Sequential choice in group settings: taking the road less traveled and less enjoyed. *Journal of Consumer Research*, **27**, 279–290.

Drolet, A. (2002) Inherent rule variability in consumer choice: changing rules for change's sake. *Journal of Consumer Research*, **29**, 293–305.

Kahn, B.E. and Isen, A.M. (1993) The influence of positive affect on variety seeking among safe, enjoyable products. *Journal of Consumer Research*, **20**, 257–270.

Kahn, B.E., Ratner, R.K. and Kahneman, D. (1997) Patterns of hedonic consumption over time. *Marketing Letters*, **8**, 85–96.

Kim, H.S. and Drolet, A. (2003) Choice and self-expression: a cultural analysis of variety-seeking. *Journal of Personality and Social Psychology*, **85**, 373–382.

Levav, J. and Zhu, R.J. (forthcoming) Seeking freedom through variety. *Journal of Consumer Research*.

McAlister, L. and Pessemier, E. (1982) Variety-seeking behavior: an interdisciplinary review. *Journal of Consumer Research*, **9**, 311–322.

Menon, S. and Kahn, B.E. (1995) The impact of context on variety-seeking in product choices. *Journal of Consumer Research*, **22**, 285–295.

Menon, S. and Kahn, B.E. (2002) Cross-category effects of induced arousal and pleasure on the internet shopping experience. *Journal of Retailing*, **78**, 31–40.

Mitchell, D.J., Kahn, B.E. and Knasko, S.C. (1995) There's something in the air: effects of congruent or incongruent ambient odor on consumer decision making. *Journal of Consumer Research*, **22**, 229–238.

Novak, D. and Mather, M. (2007) Aging and variety seeking. *Psychology and Aging*, **22**, 728–737.

Ratner, R.K. and Kahn, B.E. (2002) The impact of private vs. public consumption on variety seeking behavior. *Journal of Consumer Research*, **29**, 246–258.

Ratner, R.K., Kahn, B.E. and Kahneman, D. (1999) Choosing less-preferred experiences for the sake of variety. *Journal of Consumer Research*, **26**, 1–15.

Simonson, I. (1990) The effect of purchase quantity and timing on variety-seeking behavior. *Journal of Marketing Research*, **27**, 150–162.

family buying

James W. Gentry

Both "family" and "buying" are complex terms. For example, the question of how to define family has gained increasing relevance in postmodernity. "Family" may entail an intact nuclear family, three or four generations under one roof, a gay/lesbian couple, a single parent with children, a step-family, a single person with a network of exceptionally close ties, or even other household types. Most legal definitions are similar in substance to that of the US Census Bureau: "a family is a group of two people or more (one of whom is the householder) related by birth, marriage, or adoptions and residing together." We prefer the broader definition provided by Galvin, Bylund, and Brommel (2004, p. 6): "networks of people who share their lives over long periods of time bound by ties of marriage, blood, or commitment, legal or otherwise, who consider themselves as family and who share a significant history and anticipated future of functioning in a family relationship."

"Buying" too has complex meanings, as it involves much more than just the act of purchase. Also embedded in the term are issues of perceived need/desire for the product/service/experience; information search about the stimulus, including where to find it; transportation to the store or event, or the ordering of the item by mail, phone, or Internet; the payment for the item; its use and maintenance after purchase; its disposal (especially intergenerationally, *see* CHILDHOOD SOCIALIZATION AND INTERGENERATIONAL INFLUENCES); and the satisfaction with the memories created.

While Alderson (1957) noted that the family should be the unit of analysis in consumer research, the vast majority of family research in marketing has focused on the individual or, more precisely, on "family" decision making as the linear combination of individual preferences. Much emphasis in consumer family research has been paid to relative influence, in an effort to identify the individual most likely to make the particular decision. Measurement of "relative influence" has usually taken on an either/or perspective, using a competitive framework rather than a cooperative one more likely in the family context, and precluding investigation into constructs such as "shared influence." Thus, family consumer research has focused on individual members (Commuri and Gentry, 2000), usually the mother, rather than on interactions within the family. More recent work, however, has relied less on pencil and paper instruments and more on qualitative data that provide deeper insight into processes. An exemplar of such research was Epp (2008), who interviewed the family as a whole, as individual members, in dyads, and in triads in the context of vacation planning.

Family consumer research is an extremely dynamic phenomenon. Family gender norms are changing rapidly owing to societal changes (*see* MARKETING AND FEMINISM) such as the increasing number of households in which the wife makes a higher salary than her husband (over one-third now). Globalization is making the world smaller, and marketers are increasingly dealing with family decision making throughout the world. The most common format in the United States is the nuclear family, whereas family in the developing world is far more extended. At the same time, technology changes are having profound effects on roles within the family; for instance, consider the relative computer knowledge levels, which are generating more influence for teenagers in terms of many search processes. Family buying is a complex topic which will continue to challenge consumer researchers far into the future.

Bibliography

Alderson, W. (1957) Marketing Behavior and Executive Action, Irwin.

Commuri, S. and Gentry, J.W. (2000) Opportunities for family research in marketing. *Academy of Marketing Science Review*, **4** (5), 1–26. http://www.amsreview.org/articles/commuri08-2000.pdf

Epp, A.M. (2008) Yours, Mine, and Ours: How Families Manage Collective, Relational, and Individual Identity Goals in Consumption, (January 1, 2008), http://digitalcommons.unl.edu/dissertations/AAI3297655

Galvin, K.M., Bylund, C.L. and Brommel, B.J. (2004) *Family Communication: Cohesion and Change*, 6th edn, Allyn & Bacon, New York.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

consumer creativity

James Burroughs

Creativity is widely defined as the production of outcomes that are both novel and effective for the given context or issue. Thus, novelty and functionality form the two core dimensions of creativity. The definition of *consumer creativity* follows from the more general definition as a "departure from conventional consumption practice in a novel and functional way" (Burroughs and Mick, 2004, 403). This form of creativity could include the consumer devising a new use for an existing product, altering the form of a product to improve its performance or appearance, or combining two or more products in a unique way.

Two Forms of Creative Consumption

Instances of individuals exhibiting creativity in consumer behavior are widespread, and generally fall into one of two categories: consumer problem solving and consumer self-expression. In the first instance, the impetus for the creativity is supplied by the environment. Some type of constraint (a lack of time, money, or product availability, being the three most common; Moreau and Dahl, 2005; Bagozzi and Warshaw, 1990) prompts the individual to design his or her own consumption solution. For example, a consumer may lack a necessary ingredient for a recipe and may be forced to improvise, or a consumer who is unable to afford expensive home furnishings may use creativity to take a more affordable product and augment it to look more expensive. Many new product ideas have their origins in consumer creativity (VonHippel, 1986) (*see* CONSUMER INNOVATIVENESS), and the potential for tapping consumers as a creative resource (customer cocreation) has attracted considerable attention in recent years.

A second type of consumer creativity is more expressive in nature, such as the alteration of a car or clothing to make a statement about one's individual identity or the culture in which they are a part. Though this creativity is always embedded in a social network, it is not prompted by the environment per se, but rather comes from within the individual (Holt, 1997; Kates, 2002). Creativity is essential to every person's sense of self-worth, including feelings of autonomy, competence, and relatedness (Deci and Ryan, 2000); in modern society, these needs are increasingly met through consumption behaviors. Because this form of consumer creativity is not evoked to solve an immediate problem, the second dimension is more aesthetic than functional in nature.

There can be considerable overlap between these two forms of consumer creativity and there is no reason that an act of creativity that is highly aesthetic cannot also come in response to a very practical consumption problem. Similarly, mundane problems can be solved in very elegant ways. For this reason, it has been suggested that researchers may want to adopt a three-dimensional view of consumer creativity (i.e., novelty, functionality, and aesthetics), at least in some instances.

Antecedents of Creative Consumption

In addition to trying to understand when and why consumers may engage consumption creatively, research has also tried to understand the factors responsible for producing more creative outcomes. In other words, why are some consumers more creative than others? While a full review of all the contributing factors to creativity is not possible in this article (see Burroughs, Moreau, and Mick, 2008), four are particularly noteworthy: analogical reasoning, intrinsic motivation, domain knowledge, and risk taking.

Generally, creativity is held to flow from a confluence of interacting factors rather than any single factor. However, one factor has captured more interest and intrigue from researchers than perhaps any other, and this is analogical reasoning. Almost by definition, creativity involves taking an idea from a disparate domain and recognizing its relevance to a completely different situation. Individuals who are able to think analogically or metaphorically (as opposed to literally), consistently come up with more creative responses to consumption issues (Burroughs and Mick, 2004; Dahl and Moreau, 2002). That said, we still have very limited understanding of this fascinating cognitive process and its role in creativity.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Two less-explored but important areas of consumer creativity are domain knowledge and risk taking. The role of domain knowledge for creativity is controversial. Some advocate that a knowledge foundation is necessary for new insights to occur (Weisberg, 1999) (*see* KNOWLEDGE ACCESSIBILITY). However, others have found that too much knowledge leads to rigid thinking and is therefore detrimental to creativity. This has resulted in the inverted-U hypothesis, that creativity in an area is optimized by substantial but not overwhelming knowledge. Finally, creativity is an emergent and uncertain enterprise; thus individuals who are creative exhibit a high tolerance for ambiguity, as well as a proclivity for novelty seeking and risk taking (Sternberg and Lubart, 1996).

Another important factor is intrinsic motivation. Intrinsic motivation is sustained and intense interest in a consumption activity out of curiosity or enjoyment. Because people often engage in acts of consumption because they enjoy them (for example, a hobby), and because thinking creatively is mentally taxing, intrinsic interest is essential to the sustained effort needed to bring a potentially creative idea to fruition (Amabile, 1996).

Process Factors in Creative Consumption

Finally, a nascent area of research is trying to understand the creative process, or how creative outcomes come about. Largely based on the cognitive tradition, creativity is widely believed to have four stages: exploration, fixation, incubation, and insight (Ward, Smith, and Finke, 1999). Exploration is a preparatory stage during which many combinations of ideas and possibilities are explored (mentally and sometimes physically). Fixation is when a consumer becomes entrenched in thinking about a consumption problem or issue in one way, which often leads to mental blocks and frustration. The consumer eventually becomes fatigued and desists, at which point the process enters the incubation period. Though the consumer is no longer actively thinking about the problem, the brain continues to unconsciously contemplate it. However, the linkages that were so tightly held are now relaxed allowing new and more distal mental linkages to form. The result can be an abrupt and wholly unanticipated new approach or solution, known as the *moment of insight* (Burroughs, Moreau, and Mick, 2008). As with many other areas of creative cognition, how the brain is able to accomplish this remarkable feat is still largely a mystery.

In sum, creativity is a useful, multifaceted, and underdeveloped issue in consumer behavior. There are many opportunities for research on this topic.

See also *creativity; consumer innovativeness; knowledge accessibility; lead users*

Bibliography

Amabile, T.M. (1996) *Creativity in Context*, Westview Press, New York.

Bagozzi, R.P. and Warshaw, P.R. (1990) Trying to consume. *Journal of Consumer Research*, **17**, 127–140.

Burroughs, J.E. and Mick, D.G. (2004) Exploring antecedents and consequences of consumer creativity in a problem-solving context. *Journal of Consumer Research*, **31** (3), 402–411.

Burroughs, J.E., Moreau, C.P., and Mick, D.G. (2008) Toward a psychology of consumer creativity, in *Handbook of Consumer Psychology* (eds C.P. Haugtvedt, P.M. Herr, and F.R. Kardes), Erlbaum, New York, pp. 1011–1038.

Dahl, D.W. and Moreau, P. (2002) The influence and value of analogical thinking during new product ideation. *Journal of Marketing Research*, **39**, 47–60.

Deci, E.L. and Ryan, R.M. (2000) The what and why of goal pursuits: human needs and the self-determination of human behavior. *Psychological Inquiry*, **11** (4), 227–268.

Holt, D.B. (1997) Poststructuralist lifestyle analysis: conceptualizing the social patterning of consumption in postmodernity. *Journal of Consumer Research*, **23**, 326–350.

Kates, S.M. (2002) The protean quality of subcultural consumption: an ethnographic account of gay consumers. *Journal of Consumer Research*, **29**, 383–399.

Moreau, C.P. and Dahl, D.W. (2005) Designing the solution: the impact of constraints on consumer creativity. *Journal of Consumer Research*, **32**, 13–22.

Sternberg, R.J. and Lubart, T.I. (1996) Investing in creativity. *American Psychologist*, **51** (7), 677–688.

Ward, T.B., Smith, S.M., and Finke, R.A. (1999) Creative cognition, in *Handbook of Creativity* (ed. R. Sternberg), Cambridge University Press, New York, pp. 189–212.

Weisberg, R.W. (1999) Creativity and knowledge: abreak challenge to theories, in *Handbook of Creativity* (ed. R.J. Sternberg), Cambridge University Press, New York, pp. 226–250.

VonHippel, E. (1986) Lead users: a source of novel product concepts. *Management Science*, **32**, 791–805.

how consumers respond to price information

Kent B. Monroe

Processing of Prices as Numerical Information

We use numbers virtually every day in our lives, for example, as a form of identification of ourselves, to make telephone calls, or to pay a bill. However, the apparent ease with which we use numbers hides the fact that very complex cognitive processes are required to recognize numerical stimuli, or to make numerical comparisons or calculations. In this section, we review research results that help us understand the complexity of processing price information.

Complexity of numerical cognition processes. Number processing involves the ability to mentally manipulate sequences of words or symbols according to fixed rules. This process is of interest when a consumer is calculating or estimating the numerical difference between two prices, or adding the surcharge of shipping and handling to determine the total cost of the purchase. To compare prices, or judge the differences in prices requires that we access and manipulate a mental model of quantities similar to a mental number line. Approximation is the process by which Arabic or verbal numerals are first converted into an internal magnitude representation in our minds. This encoding is automatic, fast, and independent of which number is encoded.

When numbers are used to denote amounts, such as prices, the magnitude of the number becomes meaningful. When a price is encoded as a magnitude, either the exact value or an approximation of the exact value of the price is encoded and represented in memory. Given this encoding, we are interested in learning how people process price information when making price comparisons, when discriminating between two prices, whether the oddness of a price influences processing, and how a series of numbers that comprise a price may be remembered.

Comparing and discriminating numbers. The *distance effect* indicates that it takes longer to decide that 8 is larger than 6 than to decide that 8 is larger than 2. Moreover, the *magnitude effect* indicates that, for equal numerical distance, it is easier to discriminate small numbers (e.g., 1 vs 2) than larger numbers (e.g., 8 vs 9). The digit 5 has a special status in our numbering system. The digits 2, 3, and 4 are initially encoded as small, while the digits 6, 7, and 8 are encoded as large. That is, the mental representation of numbers below 10 is divided into numbers above and below 5. Thus, in a price comparison, task people first automatically encode the two prices independently and then classify the prices as "small" or "large" (Monroe and Lee, 1999).

When judging which number in a pair is *larger*, people find the task easier when both numbers are large than when both are small. The opposite is true when people have to decide which item is smaller. It has been shown that when people are trying to determine which of the two numerals is larger, it is an easier task if the larger of the compared numerals is displayed in larger font size. Similar results occur when people are trying to determine the smaller of compared numerals and the smaller numeral is displayed in smaller font size (Coulter and Coulter, 2005).

As suggested above, some processing of price information may be more automatic than others. Also, price comparisons in the lower price range (i.e., smaller numbers involving fewer digits), which are more typical for most supermarket purchases, may be more likely to be processed automatically. Moreover, identifying whether a specific price is higher than another price will be processed faster when that price is indeed higher than the comparison or reference price.

The left digit and price endings effects. Research has demonstrated that under a variety of different tasks and judgments, even digits are processed faster and/or more accurately than odd digits. Possible reasons for this odd price effect on consumer behavior can be divided into effects of *price perception* (how people encode price information into their minds) and effects of *price cognition* (how people process and interpret the price information).

Consumers have a tendency to underestimate 99¢-ending prices. Given that consumers seem to have more difficulty recalling odd-ending prices than round prices (e.g., $300) could

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

imply that odd-ending prices are not encoded completely or that $300.00 is easier to encode as a magnitude than $299.99. One reason is prices ending in 99¢ are more difficult to process and as a result people may encode an approximation of the price rather than the precise price. Also, people tend to misperceive the difference between two prices when the left-most digits differ, odd price endings are used, and computations are necessary to determine the magnitude of difference (Manning and Sprott, 2009; Thomas and Morwitz, 2009). Knowing that perceived price has *meaning* that differs from its objective meaning in terms of perceived monetary sacrifice suggests that using inappropriate price endings could have unwanted implications for sellers.

Thus far, we have considered only a single price for a product or service, what are called *unidimensional prices*. However, offers often are provided with multiple price attributes. For example, an automobile may be advertised with a price for the vehicle, along with a combination of monthly payments, amount of down payment, and effective interest rate. Catalog and Internet offers typically quote selling prices and shipping and handling charges. When there are multidimensional prices, additional computations are required to evaluate the offer, some of which can be difficult. Also, some retail price promotions provide multiple discounts, for example, 40% off all merchandise with an extra 15% off if purchased at a certain time. When odd price endings are used, or when an item has a complicated price, for example, $23 977, these types of offers affect consumers' abilities to evaluate prices and to know exactly what they are paying for the item. The computational difficulty associated with multidimensional pricing is influenced by the price ending used, the perceived numerical distance between comparative prices, as well as the calculations necessary to determine the price to be paid.

Remembering prices. Previous research has attempted to determine the extent that consumers do remember prices they have paid assuming that consumers are *consciously* involved in their purchasing behaviors and that product evaluations and choice decisions are a function of what information is accessible in memory. What is accessible in memory often is measured by what consumers *can consciously remember*. The notion is that consumers frequently make judgments that a particular item is "too expensive" or "a real bargain" based on some prices that they recall from past shopping experiences, and these recalled prices form a basis on which the consumers' reference price is formed.

However, research measuring consumers' ability to remember prices of recently purchased items report that a relatively low proportion of consumers can accurately recall prices of recently purchased products (Monroe and Lee, 1999). Such research evidence could lead to the conclusion that consumers do not consciously attend to price information when either considering or actually making purchase decisions. Nevertheless, a shopper who does not recall the price of, say, the box of snack bars she just put into her cart may tell the interviewer that it is expensive. Thus, consumers may encode price information into memory in different representational forms.

Consumers may process and retrieve price information either consciously or nonconsciously. When price information is processed consciously, consumers pay attention to the price, make judgments regarding the value of the product using information that is either present in the external environment or retrieved from memory, and finally make a purchase decision. When consumers consciously process the actual price information, a magnitude representation of the price and their evaluative judgment may be transferred from working memory into long-term memory (Vanhuele, Laurent, and Dreze, 2006).

Alternatively, it is possible that only their comparative or evaluative judgments, and not the actual price information, are transferred into long-term memory. In this case, the consumers would not be able to recall the actual price when asked to do so at a later point in time. This scenario is consistent with the idea that an evaluation of a product may be more easily remembered than the price when the evaluation is formed at the time of exposure (Adaval and Monroe, 2002). When price information is processed nonconsciously, the consumer does not pay particular attention to the prices. Nonetheless, a judgment regarding the value of the product and a purchase decision may have been made. When

nonconscious processing of price information occurs, the consumer may not be able to recall the price of the product at a later time, but she may still be able to indicate that the product is "too expensive," "a bargain," or "reasonably priced," indicating that the price information has been processed and evaluated. Thus, consumers' price comparisons and evaluations are affected by price information to which they have been previously exposed, even though they may not consciously remember the actual prices relevant to their judgments.

Reference Prices

Price judgments are comparative in nature. That is, for consumers to determine that a price is acceptable, too high, or low, that price must be compared to another price, whether that other price is in memory or available nearby. The other price serves as a comparison standard or reference price. Researchers have proposed and provided evidence for the existence of reference prices (Mazumdar, Raj, and Singh, 2005). These propositions have been based on adaptation-level theory and assimilation-contrast theory (social judgment theory).

Adaptation-level theory. Adaptation-level theory assumes that stimuli are judged with respect to internal norms (*adaptation levels*) representing the combined effects of present and past experiences. For any consumer, the adaptation-level price for a specific product category is a function of the frequency distribution of prices for that category, the relative magnitude of the prices, the range of prices, and the dispersion of prices from the average price. It is influenced by past experience and by the sequence that the consumer observes the prices for the category. The concept of adaptation level or reference price supports the important point that *consumers judge or evaluate prices comparatively*; that is, the acceptability of a price is judged by comparison to another price.

When considering how consumers perceive prices, it has to be borne in mind that perception is relative. In other words, a specific price is compared to another price, or a reference price. When sellers advertise both the offered price and a (higher) comparative regular price, they are suggesting a reference price for consumers' comparisons. To convince consumers to use the higher price as a reference price, sellers may include such words as "formerly", "regularly," and "usually," to describe the higher price. The judgment of acceptability depends not only on consumers' price expectations but also on information provided in promotions or advertisements.

Assimilation-contrast theory. Similarly, the basic tenet of assimilation-contrast theory is that new stimuli encountered by an individual are compared against a background of experience with the stimuli category. This experience forms an individual's reference scale that serves as a basis for the individual to compare and evaluate other stimuli *perceived as related to* the reference stimuli.

If a new price is encountered for evaluation, it may also serve as an anchor in that it may result in a changing of the reference scale. Both adaptation-level theory and assimilation-contrast theory suggest that the reference price range should move in the direction of this new price. If the new price is relatively high, the reference price range moves toward the new price and the previously lower prices become further away from the reference price scale. If judgments of these previously perceived low prices in the category do not change when the new price is encountered, an *assimilation effect* is said to occur. The assimilation effect occurs because the new price is perceived as similar to the reference prices. However, if the new price moves the reference price range sufficiently higher, the original low prices will be perceived as lower than previously, and a *contrast effect* is said to occur.

The prevailing range of prices for a product category affects the consumer's reference price for that category. Since the width of the price range is affected by the lowest and highest price (end prices) in the range, these two end prices also affect price judgments. Thus, there are three different price cues affecting price judgments: the reference price, the lowest price, and the highest price in the range of prices available for judgment. Prices used by consumers to judge other prices (reference price, end prices) are called *anchoring stimuli*. By *price judgment* we

mean a consumer's assessment of whether a price is too low, just right (acceptable), or too high.

In summary, a reference price is any price in relation to which other prices are perceived. These two theories clearly suggest that a reference "price" is actually an internally held judgment scale concerning prices for a product-price category.

Price Thresholds

Absolute price thresholds. An important concept introduced in our discussion of reference price was the *acceptable price range*. This concept implies that a consumer has a lower price threshold and that there are prices greater than $0, which are unacceptable because they are considered to be too low, perhaps because consumers are suspicious of the product's quality. Also, it is recognized that at specific points in time there is a maximum price that consumers are willing to pay for a product or service (upper price threshold). The important point is that there is not just one acceptable price for a product or service; instead, there is some range of acceptable prices.

Variations in the level and width of consumers' acceptable price ranges are influenced by a number of factors. For a specific product category, the upper acceptable price threshold is lower if consumers perceive that there are similar alternative offerings available. However, if customer satisfaction increases or consumers become more loyal, then the upper threshold tends to be higher. Conversely, if customer satisfaction declines leading to lower consumer loyalty, consumers' upper price threshold would become lower. If consumers are not knowledgeable about prices for a product category, their acceptable price ranges tend to be relatively narrow and their lower and upper acceptable price limits tend to be lower than more knowledgeable consumers. Consumers who infer quality on the basis of price tend to have higher acceptable price levels, higher upper acceptable price limits, and wider acceptable price ranges. Finally, consumers who are consciously concerned about prices tend to have lower acceptable price levels, lower upper acceptable price limits, and narrower acceptable price ranges (Monroe, 2003).

Differential price thresholds. Usually a consumer has alternative choices available for a purchase and selects from among these choices. The prices of these alternative choices may provide cues that facilitate the decision process. However, even if the numerical prices are different, it cannot be assumed that the prices are *perceived* to be different. Hence, the problem becomes one of determining the effect of *perceived price differences* on consumer choice.

The perception of a price change or difference depends on the magnitude of the change or difference. Also, people are more sensitive to perceived price increases than to decreases. The immediate implication is that consumers will be more sensitive to price changes for some products, that is, have lower differential price thresholds. But, for other products, a similar price change may not be perceived. The differential price issue is how the price of one product is perceived to differ from the price of another offering that consumers believe is an alternative choice to consider. These alternative products could be sold by different sellers competitively or they could be alternative models of the product sold by a single seller.

Price and Consumers' Perceptions of Value

Behavioral research has provided explanations of how people form value judgments and make decisions when they do not have perfect information about alternatives. These findings further our understanding of why consumers may be more sensitive to price increases than to price decreases, and how they respond to comparative price advertisements (e.g., regular price $65, sale price $49), coupons, rebates, and other price promotions.

In this section, we return to the idea that people seldom are good information processors and that they often take shortcuts (decision heuristics). These shortcuts may lead to errors in judgment and choice, but they may also facilitate the choice process. Several of these decision heuristics can help us understand how price influences perceptions of value and eventual product choice. First, the *context* of the purchase decision, including the way the offer is presented, *frames* the consumer's evaluation and

choice. For example, the way a sale is advertised will influence consumers' judgments about the value of the offer. The context of the place of purchase affects a consumer's internal reference price or expectations about prices in that place.

Another important point is the *anchoring* effect mentioned earlier. As we observed earlier, the order of price presentation anchors consumers' judgments as do the low or high prices in a product line. Also, people tend to adapt to prices that are presented as the original prices. For example, antique dealers often overprice their items, anticipating that consumers will want to negotiate over the price. The initial high price serves as an anchor, and generally the negotiated price is higher than it would have been without this initial high anchor price. Prices that are irrelevant to the considered purchase can anchor judgments about what is acceptable to pay for an item (Nunes and Boatwright, 2004).

Perceived value represents a trade-off between consumers' perceptions of quality and sacrifice. The degree that consumers believe there is a price–quality relationship influences their value perceptions and willingness to buy. When comparing prices, consumers' judgments are influenced by the relative or perceived differences between the actual or offer price and the reference price. Using these points about price perception, we now show how perceived price influences consumers' judgments of value.

The price–perceived quality relationship. Consumers do not use price solely as a measure of cost (sacrifice). They also use price as an indicator of product quality (Rao and Monroe, 1989). Consumers are assumed to assess product or service quality by the use of cues. Products, services, or stores can be conceptualized as consisting of an array of cues that may serve as indicators of quality. Consumers may use these cues if the cues help them predict the quality of the product or service and when they have confidence that they can use and judge the cues accurately. *Extrinsic cues* are product-related attributes – price, brand name, packaging, but they are not part of the product. *Intrinsic cues* are also product-related attributes, but they cannot be changed without altering the physical properties of the product. Consumers rely on extrinsic cues and intrinsic cues when evaluating quality.

When consumers become familiar with a product they are more likely to use intrinsic cues rather than price or other external cues as indicators of product quality. However, highly familiar consumers (experts) use either price or intrinsic cues as indicators of quality, depending on whether their knowledge includes information about the reliability of price as a quality indicator. That is, if consumers know that there is a positive price–quality relationship in the market, they will probably use price as a quality indicator. It has been argued further that if consumers know that there is a weak price–quality relationship in the product market, they will be more likely to use intrinsic product cues to assess product quality. Thus, the strength of the use of price or other external cues, such as brand or store name, as indicators of product quality, depends on the relative perceived differences between different cues and the degree that consumers know about the product and actual price–quality relationships.

The price–perceived monetary sacrifice relationship. A consumer's perceptions of the sacrifice, "give up," or loss incurred by paying the monetary price for a product may vary according to a variety of situations and conditions. For example, does a consumer perceive paying $90 for a dental cleaning to be an equivalent sacrifice as paying $90 for a ticket to a concert or sporting event? Indeed consumers may perceive equivalent monetary prices as representing very different sacrifices.

For some products or services, consumers would prefer not to make the expenditure. Thus the perceived sacrifice of paying a specific price for some products is psychologically more "painful" than others, even though the monetary outlay may be equivalent. Thus, the more the consumers are reluctant or hesitant to spend money on certain products, the more likely are they to search for bargains or lower prices, and they will be more sensitive to price changes and to price differentials between alternative choices.

If consumers believe that sellers have increased prices to take advantage of an increase in demand, or a scarcity of supply, without a corresponding increase in costs, then such price increases would be perceived to be unfair (Bolton, Warlop, and Alba, 2003; Xia, Monroe,

and Cox, 2004). In situations where one category of consumers receive the benefit of a lower price for an equivalent product or service but another category of consumers do not, the price-disadvantaged consumers may perceive that the price they pay is unfair (Haws and Bearden, 2006). Such situations typically occur when sellers provide discounts for consumers with certain characteristics, for example, age or preference status. Further, if there is no perceived discrepancy between the effort that these favored consumers make to qualify for the lower price relative to the disadvantaged consumers, then they will be receiving the same benefits as the disadvantaged consumers but incurring a smaller monetary sacrifice. Similarly, if consumers perceive that the prices they pay are higher than their comparable reference group, then such prices are likely to be perceived to be unfair.

Thus, a price judged as unfair can lead to lower perceptions of value and a reduction in willingness to pay; that is, a perceived disadvantaged price inequity, or a loss, increases consumers' perceptions of sacrifice, thereby decreasing their perceptions of value and willingness to buy.

Integrating Price, Quality, and Sacrifice Information

When consumers use price as an indicator of cost or sacrifice, increasing price has the effect of reducing perceived value for the product or service. On the other hand, if consumers use price as an indicator of quality or benefits, increasing price has the effect of increasing perceived value. We know that consumers generally are not able to assess perfectly product or service quality (the ability of the product to provide satisfaction). We also know that perceptions of monetary sacrifice for the same specific price may vary according to whether the price is perceived to be fair, the context in which the price is presented, the consumers' acceptable price range, and the extent that consumers are concerned about price. We also understand that perceived value represents a mental trade-off between consumers' perceptions of quality and sacrifice and is positive when perceptions of quality are greater than their perceptions of sacrifice.

A very important issue is how consumers are able to integrate these different bits of information to determine their overall perceptions of value. In other words, information about the product or service attributes and benefits represents positive information (i.e., what the consumer gains from acquiring the product or service). Considering price as an indicator of cost represents negative information (i.e., what the consumer gives up when acquiring the product or service). Attempting to integrate positive and negative information simultaneously to determine an overall value judgment is a difficult mental task.

Although we have not been able to clarify the exact nature of this trade-off between perceived quality and sacrifice, there are some important implications given what we do know. The weight consumers attach to quality and sacrifice depends on the relative magnitude of the product's price. Generally, both quality and sacrifice are perceived to be low when price is perceived to be relatively low, and to be high when price is perceived to be relatively high. However, when evaluating a low-priced product, consumers appear to weigh quality more heavily than sacrifice, thus judging the product as low in value. Yet, when judging a high-priced product, consumers tend to weigh sacrifice more heavily than quality, again judging the product as low in value. These generalizations assume that consumers have sufficient motivation and cognitive resources to process the price and other attribute information to determine product value (Suri and Monroe, 2003). However, these assumptions about motivation and available cognitive resources are not always correct.

Decomposing Perceived Product Value

The overall perceived value of a product being evaluated is its (i) *acquisition value* (the expected benefit to be gained from acquiring the product less the net displeasure of paying for it) and (ii) *transaction value* (the perceived merits or fairness of the offer or deal) (Grewal, Monroe, and Krishnan, 1998).

Acquisition value. Consumers' perceptions of acquisition value represent a cognitive trade-off between the benefits they perceive in the product

and the sacrifice they perceive to be required to acquire the product or service by paying the monetary price of the product. In part, the perceived benefits of a product are related to the consumers' judgments about the product's quality. Lacking perfect information about the inherent quality of the product, many consumers tend to believe that there is a positive relationship between a product's price and its quality ("You get what you pay for"). Thus, other things remaining the same, a higher priced product would be perceived to provide more benefits because of its higher perceived quality. However, at the same time, a higher price increases consumers' perceptions of their sacrifice. Thus, within some range of prices, the perceived benefits in the product will be larger than the perceived sacrifice, and consumers will perceive that there is positive acquisition value in the product. Generally, the greater the perceived acquisition value, the greater is the likelihood that consumers would be willing to purchase the product. However, besides evaluating the product's value, consumers also evaluate the offer itself.

Transaction value. *Transaction value* is defined as the consumers' perceived merits of the offer or deal. Of concern here is how consumers evaluate a purchase situation in which the consumer gains a product but gives up the money paid for the product. Consumers first *judge* the value of the offer and then *decide* whether to make a purchase. To explain the role of price in this process, three price concepts are used. The *perceived benefit* of the product is equivalent to the utility inherent in the *maximum acceptable price* the consumer would be willing to pay. The *acquisition value* of the product is the perceived benefits of the product at this maximum price compared to the actual selling price, that is, $p_{\max} - p_{\text{actual}}$. The *transaction value*, or the perceived merit of paying the actual price, is determined by comparing the consumer's reference price to the actual price, that is, $p_{\text{ref}} - p_{\text{actual}}$. Transaction value is positive if the actual price is less than the consumer's reference price, zero if they are equal, and negative otherwise.

Researchers have discovered that when transaction value is present it enhances acquisition value, but does not directly influence consumer behavior (Grewal, Monroe, and Krishnan, 1998). Thus, acquisition value is determined by the consumers' perceptions of quality or benefits to be received plus perceived transaction value, which represents the comparison of the selling price to the consumers' reference price. One important implication of the finding that perceived transaction value enhances consumers' perceived acquisition value is consumers need to feel confident that they either can determine quality prior to purchase, or they can infer quality because the various signals of quality used by the sellers are appropriate indicators of quality.

Summary

Throughout this article, we have developed a behavioral explanation of how consumers perceive price and how these perceptions influence their perceptions of value. Adaptation-level theory indicates not only that there is a reference price but also that it changes. Reference price is affected by contextual effects such as frequency of previous price changes, consumers' expectations about future prices, the order that price information is presented to consumers, the advertisement of prices, and the intensity of price promotion.

Assimilation-contrast theory indicates that there is a range of acceptable prices. This theory also suggests that the acceptable price range may be affected by the amount of price variation for a product category as perceived by consumers. Another implication derived from assimilation-contrast theory is that there is likely to be a range of prices around the reference price within which little change in demand is likely in response to a price change, that is, the price resulting from the price change may not be perceived as very different.

Finally, the psychological argument recognizes that people respond differently to perceived gains and perceived losses and suggests that consumers are more sensitive to price increases (perceived loss) than to price decreases (perceived gain). However, once the price difference (i.e., $p_{\text{ref}} - p$) is perceived to be important by the consumer, there is likely to be a more noticeable change in demand. These price differences can occur because of either a price

increase or decrease, or because of comparative price advertising in which the seller provides an external reference price for comparison.

See also *brand value; consumer memory processes; customer analysis; demand elasticity; pricing strategy*

Bibliography

Adaval, R. and Monroe, K.B. (2002) Automatic construction and use of contextual information for product and price evaluations. *Journal of Consumer Research*, **28**, 572–587.

Bolton, L.E., Warlop, L., and Alba, J.W. (2003) Consumer perceptions of price (un)fairness. *Journal of Consumer Research*, **29**, 474–491.

Coulter, K.S. and Coulter, R.A. (2005) Size does matter: the effects of magnitude representation congruency on price perceptions and purchase likelihood. *Journal of Consumer Psychology*, **15**, 64–76.

Grewal, D., Monroe, K.B., and Krishnan, R. (1998) The effects of price-comparison advertising on consumers' perceptions of acquisition value, transaction value, and behavioral intentions. *Journal of Marketing*, **62**, 46–59.

Haws, K.L. and Bearden, W.O. (2006) Dynamic pricing and consumer fairness perceptions. *Journal of Consumer Research*, **33**, 304–311.

Manning, K.G. and Sprott, D.E. (2009) Price endings, left-digit effects, and choice. *Journal of Consumer Research*, **36**, 328–335.

Mazumdar, T., Raj, S.P., and Singh, I. (2005) Reference price research: review and propositions. *Journal of Marketing*, **69**, 84–102.

Monroe, K.B. (2003) *Pricing: Making Profitable Decisions*, McGraw-Hill/Irwin, Burr Ridge.

Monroe, K.B. and Lee, A.Y. (1999) Remembering versus knowing: issues in consumers' processing of price information. *Journal of the Academy of Marketing Science*, **27**, 207–225.

Nunes, J.C. and Boatwright, P. (2004) Incidental prices and their effect on willingness to pay. *Journal of Marketing Research*, **41**, 457–466.

Rao, A.R. and Monroe, K.B. (1989) The effect of price, brand name and store name on consumers' perceptions of product quality: an integrative review. *Journal of Marketing Research*, **26**, 351–357.

Suri, R. and Monroe, K.B. (2003) The effects of time constraints on consumers' judgments of prices and products. *Journal of Consumer Research*, **30**, 92–104.

Thomas, M. and Morwitz, V.G. (2009) Heuristics in numerical cognition: implications for pricing, in *Handbook of Pricing Research in Marketing* (ed. V. Rao), Edward Elgar Publishing, pp. 132–149.

Vanhuele, M., Laurent, G., and Dreze, X. (2006) Consumers' immediate memory for prices. *Journal of Consumer Research*, **33**, 163–172.

Xia, L., Monroe, K.B., and Cox, J.L. (2004) The price is unfair! A conceptual framework of price fairness perceptions. *Journal of Marketing*, **68**, 1–15.

consumer behavior analysis

Gordon R. Foxall

Introduction

Although the main theoretical perspective of consumer research has been cognitive, behaviorism receives intermittent mention as a possible contributor, though its potential for understanding consumer choice remains underdeveloped. Some consumer theories incorporate it to deal with routine aspects of consumer behavior, while other treatments approach it as a source of marketer, especially retailer, tactics. The possibility that behaviorism provides insights into the explanation of consumer choice has been neglected. *Consumer behavior analysis* attempts to redress the balance by exploring the nature of behaviorist explanation and its capacity to enlighten consumer research.

Especially in terms of the theoretical and empirical advances made by behavior analysts in the field of verbal behavior and behavioral economics during the last two decades, behaviorism promises to extend the investigation and explanation of consumer behavior beyond the limits of a purely cognitive approach. Behavioral economics, in particular, combines the rigor of operant theory with the methods of experimental economics with the aim of understanding aspects of consumer choice that cognitive consumer psychology has often neglected, such as gambling, addiction, and health-related behaviors as well as more routine features of consumer behavior such as product and brand choice. Consumer behavior analysis adds to this the contextual framework of consumer decision making in marketing-oriented economies, adding further to the interdisciplinary base of the psychological investigation of economic behavior.

The Behavioral Perspective Model

The original aim of the research program was to ascertain whether a model of consumer choice based on a radical behaviorist framework was feasible, and, if so, what the epistemological nature of such a route to explanation would be (Foxall, 2004a). So the aim has never been to change the paradigm in consumer research in favor of behaviorism, but to test a radical behaviorist depiction of consumer choice to its limit, adding in other approaches to explanation only as and when they became essential. The earliest stages involved critiques of the central explanatory devices assumed in consumer research at the time, notably the ideas that attitudes and intentions inevitably precede, prefigure, and determine consumer behavior, or that novel behavior on the part of consumers was explicable in terms of underlying traits of "innovativeness." The development of the sought-after model of consumer behavior based on radical behaviorism resulted in the behavioral perspective model (BPM), which is an elaboration of the "three-term contingency," the basic explanatory device of operant psychology in which a discriminative stimulus (S^D) marks the occasion on which a particular response (R) is likely to be rewarded or reinforced (S^R), that is, to increase in frequency, or punished (S^P), that is, to decrease. The three-term contingency is usually depicted as

$$S^D : R \longrightarrow S^{R/P}$$

where : indicates that the probability of an operant response is increased in the presence of the S^D, while by definition that response leads to consequences that are reinforcing and/or punishing. The model adapts these key elements to suit the interpretation and prediction of human economic behavior in affluent societies. The BPM is shown in Figure 1.

Consumer choice is the outcome of the consumer's learning history meeting the current consumer behavior setting, the point at which the experience of consumption meets a new opportunity to consume. This intersection is the *consumer situation*, the immediate determinant of approach/avoidance responses involved in purchase and consumption. The *consumer behavior setting* comprises the stimulus antecedents of that behavior, some of which would have been present on earlier consumption occasions. Given the individual's learning history, that is, past choices and the reinforcing/punishing consequences they have had, these initially neutral stimuli are transformed into the discriminative stimuli

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

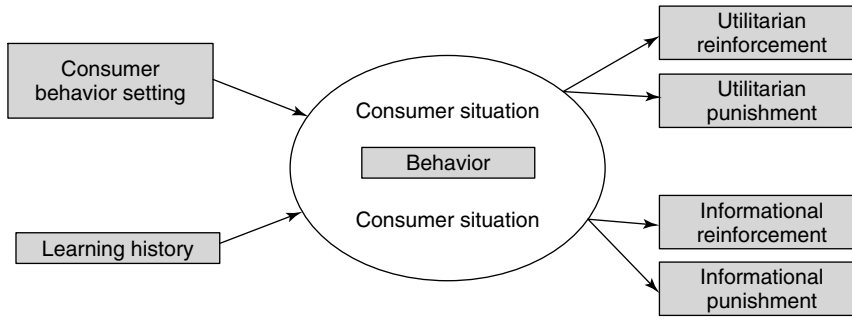


Figure 1 Summative behavioral perspective model.

that set the occasion for current choice; in particular, the individual's consumption history invests them with meaning, in the sense of a capacity to generate specific kinds of approach and/or avoidance behaviors, which produce consequences that regulate the rate of recurrence of the behaviors that produced them. The consumer situation consists also of motivating operations (MOs) such as rules that invest the consequences inherent in the discriminative stimuli with additional motivating or inhibitory power by making the consequences of radical behaviorism appear more or less reinforcing, more or less punishing.

The consumer behavior setting. Like the three-term contingency, the BPM specifies behaviorally antecedent stimulus conditions (the behavior setting) but combines the concepts of discriminative stimuli and motivating operations by means of the construct of behavior-setting *scope*, the extent to which these setting elements encourage or inhibit the behavior predicted to occur in such settings. Settings of purchase and consumption are all *relatively* open, but differ from one another along a restricted continuum of closed–open consumer behavior settings. Waiting in line at the bank to pay in a check occurs in a relatively closed consumer behavior setting: there is probably no alternative to being there and waiting until a teller becomes available, standing in an orderly fashion is encouraged both by the physical style of the building and by the social arrangements, deviation from the established behavior program of the setting is likely to be punished by stares or glares. An open consumer behavior setting encourages a wider range of alternative behaviors. In a bar, for instance, all types of beverages and snacks may be available, there may be TV to watch, talking loudly may not be discouraged, and even singing and dancing may be possible. The customer is free to leave at any time, even if only to go to another bar in the vicinity – at least far freer than he or she would be to leave the bank and find another at which to present the check.

Patterns of reinforcement. The consequences of economic behavior fall into three types: utilitarian reinforcement, which consists in the functional outcomes of behavior, informational reinforcement, which stems from the symbolic outcomes, principally performance feedback, aversive/punishing consequences, and the costs of purchase and consumption. Such aversive outcomes can themselves be subdivided into those that are utilitarian in nature and those that are symbolic. Utilitarian reinforcement consists in the direct usable, economic, and technical benefits of owning and consuming a product or service, while informational reinforcement inheres in benefits of ownership and consumption, which are usually social in nature and consist in the prestige or status as well as the self-esteem generated by ownership and consumption. The driver of the cheapest car available is principally concerned with the utilitarian benefits that all cars provide: most obviously, door-to-door transportation. Informational reinforcement, on the other hand, is more likely to involve a lifestyle statement by which the consumer seeks to convey his or her social status or to bolster esteem and/or reported feelings of self-esteem. The driver of a Porsche

not only clearly gets from A to B in it but also receives social esteem and status from others as well as personally conferred self-esteem. These constitute symbolic rewards of consumption. Most products have an element of both the instrumental and the symbolic. A mobile phone not only provides communications services when and where the consumer wants them; because it has interchangeable colored cases, it may also signal to that consumer's social group that he or she is "cool" (or, a year later, "not so cool").

Patterns of reinforcement, based on combining high and low levels of utilitarian reinforcement and high and low levels of informational reinforcement, suggest four operant classes of consumer behavior (Figure 2). The four broad classes of consumer behavior can be inferred as follows. (i) *Accomplishment* is consumer behavior reflecting social and economic achievement: acquisition and conspicuous consumption of status goods, displaying products and services that signal personal attainment. (ii) *Hedonism* includes such activities as the consumption of popular entertainment. (iii) *Accumulation* includes the consumer behaviors involved in certain kinds of saving, collecting, and installment buying. (iv) *Maintenance* consists of activities necessary for the consumer's physical survival and welfare (e.g., food) and the fulfillment of the minimal obligations entailed in membership of a social system (e.g., paying taxes). Both types of reinforcers figure to differing extents in each of the four classes. Adding in the scope of the current behavior setting leads to the eightfold way depicted in Figure 3, which shows the variety of contingency categories that exclusively constitute a functional analysis of consumer behavior. We take a closer look at the four broad operant classes of consumer behavior with the added complexity of consumer behavior-setting scope added in.

Accomplishment. In an open setting, this consists, in general, in the *purchase and consumption of status goods*: luxuries and radical innovations such as exotic vacations and iPhones. The items in question are possessed and used for the pleasure or ease of living they confer, the well-being they make possible for the individual; they thereby provide extensive hedonic rewards. As status symbols, their conspicuous consumption strengthens the behavior in question. They attest directly, and often publicly and unambiguously, to the consumer's attainments, especially economic. Goods in this category are usually highly differentiated – by novel function in the case of innovations, by branding in the case of luxuries. In a closed setting, accomplishment can be generally described as *fulfillment*, personal attainments gained through leisure, often with a strong element of recreation or excitement as well as achievement. It might, for instance, include both the completion of a personal development seminar and gambling in a casino, both of which are maintained by high levels of utilitarian and informational consequence in fairly closed settings.

Hedonism. In an open setting, this consists of *popular entertainment*: viewing TV game shows, which provide near-constant utilitarian reward, or reading mass fiction, which contains a sensation on almost every page. iPods and, in their day, DVDs have made such reinforcement more immediate to the point of ubiquity. Mass culture presents frequent and predictable, relatively strong and continuous hedonic rewards, which are not contingent on long periods of concentrated effort. The arrangement of reinforcers is such that viewing,

	High utilitarian reinforcement	Low utilitarian reinforcement
High informational reinforcement	Accomplishment	Accumulation
Low informational reinforcement	Hedonism	Maintenance

Figure 2 Operant classes of consumer behavior defined by pattern of reinforcement.

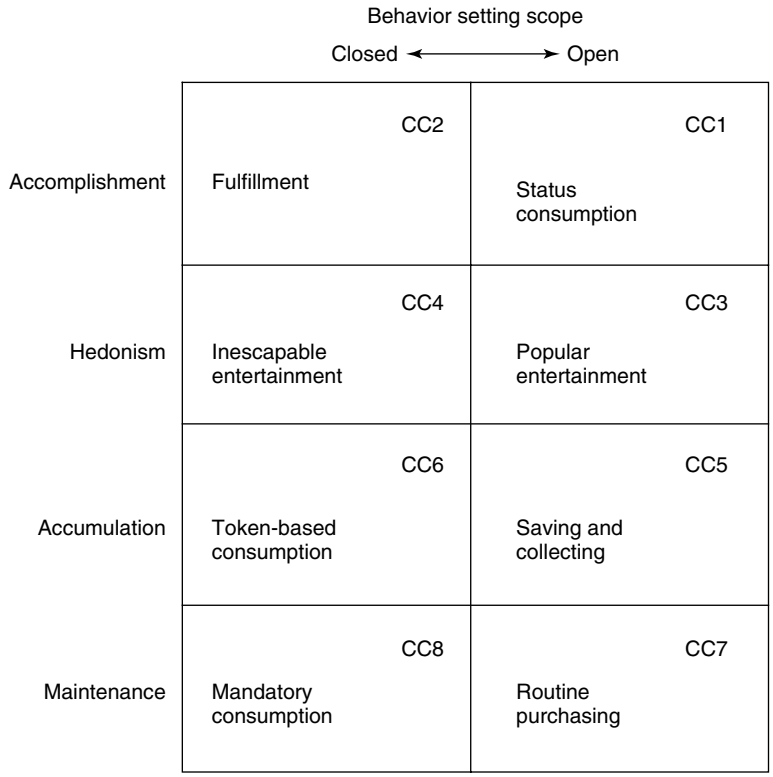


Figure 3 BPM contingency matrix.

listening, or reading for even a short interval is likely to be rewarded. Informational feedback is more obvious on some occasions than others, as when game shows allow the audience to pit their own performances against that of the competing participants, but it is not the main source of reward. Hedonism in closed settings often consists of inescapable entertainment and amelioration. The behaviors in question are potentially pleasurable but may become irksome because they are unavoidable. Consumption of these products and services may be passive rather than active. An example arises when long-haul airline passengers purchase meals and movies along with their travel. The meals are usually consumed, like the in-flight movies, which follow them, without alternative. The setting, which cannot be other than highly restrictive for safety reasons, is further closed by the pulling of blinds, the disappearance of cabin staff, the impossibility of moving around the plane, and the attention of fellow passengers to the movie. To try to read or engage in other activities may even invite censure.

Accumulation. In an open setting, accumulation is *saving and collecting*: purchases for which payments are made prior to consumption – installments for a holiday, which can only be taken once the full amount has been paid. Discretionary saving with the intention of making a large purchase once a certain amount has accumulated is also included as are promotional deals requiring the accumulation of coupons or other tokens before a product or service can be obtained. The immediate reward is informational, feedback on how much one has accumulated, how close one is to the ultimate reinforcer. Accumulation occurring in a closed setting may be described, in general terms, as *token-based buying* which also involves

collecting – via schemes in which payment for one item provides tokens with which future rewards are obtained. The points earned by frequent flyers constitute informational reinforcers; some hotels offer gifts to frequent customers who accumulate points on each stay. The setting is relatively closed, because once one becomes a member of the scheme there is an incentive to remain in it.

Maintenance. In an open setting, maintenance is *routine purchasing and consumption* including buying goods necessary for survival as in the habitual purchasing of groceries at a supermarket. Comparatively, few such consumers are brand loyal in the sense of always choosing the identical brand in a long sequence of shopping trips. There is so much choice that the consumer enjoys considerable discretion among versions of the product. Maintenance in closed settings is *mandatory purchase and consumption* and includes consumer behaviors necessary to remain a citizen: the payment of taxes for public goods, payments into pension schemes, payments (as in the United Kingdom) of TV licenses.

Research

Early applications of the BPM were as an interpretive device; issues central to consumer research were examined not within the usual cognitive framework but from a behavioral perspective: the relationship between consumers' attitudes and behavior, the adoption of innovations, spending, consuming and saving, "green" consumption, and the nature and influence of marketing management were interpreted within the new framework (Foxall, 2009). Most of this work was conducted within an ambience of testing out behavior analysis by comparing it with the conventional information processing accounts of consumer choice, a desire to promote a clash of competing explanations. Four strands of later research exemplify the empirical testing of the model: the prediction of consumers' emotional responses to retail and consumption environments, the application of matching and maximization techniques to consumer choice, the analysis of consumer demand (behavioral economics), and experimentation.

Emotional response. The first aim was to test the model as a whole, to understand whether (and if so how) the three structural variables of setting scope, utilitarian reinforcement, and informational reinforcement interacted. Eight empirical studies conducted in a variety of cultural settings employed a tripartite classification to emotion based on Mehrabian and Russell's theory of environmental psychology in which the central emotions are pleasure, arousal, and dominance. Measures of these emotions were hypothesized as predictable in consumer contexts defined by the BPM, so that utilitarian reinforcement was expected to generate pleasure, which implied satisfaction and utility; informational reinforcement to be associated with arousal, which reflects behavioral feedback; and the scope of the consumer behavior setting would correspond to dominance, i.e., being in control. The results of this work confirm the predictions (Foxall, 2005). The expectations of a larger pleasure score for higher utilitarian reinforcement; similarly, a larger arousal score was found for greater informational reinforcement; and a larger dominance score characterized a more open consumer behavior setting scope. The results indicate that by using the pleasure, arousal, and dominance measures as predicted verbal responses to the consumer situations defined by the BPM contingency matrix, it is possible to make useful predictions of consumer behavior.

Matching and maximization. *Matching* is the tendency of animals and humans to distribute their responses between two choices in proportion to the patterns of reward programmed to be contingent on each choice. The key dependent variable is not the single response that needed contextual explication in terms of a single contingent reinforcer: it was the relative frequency of responding, which he or she explained by reference to the relative rate of reinforcement obtained from the behavior. Animals with two opportunities to respond (pecking key A or key B), each of which delivers reinforcers (food pellets) on its own reinforcement schedule, allocate their responses on A and B in proportion to the rates of reward they obtain from A and B. This phenomenon has been replicated in numerous species including humans

and has found applications in behavior modification and organizational behavior management, to name but two relevant fields. In particular, it provides a framework for the behavioral analysis of consumption. The phenomenon is particularly well researched in contexts that require an individual to allocate a limited period between two choices, each scheduled to produce reward at a different rate.

Most choices for human consumers require the allocation of a fixed income between alternative choices, each of which exacts a different monetary sacrifice. In this case, responses take the form of surrendering money in varying amounts, while the reward is the receipt of a fixed amount of the good in question. Price is the ratio of units of money that must be exchanged for units of the good. Both matching and maximizing theories make a similar prediction of behavior on such schedules: the individual will maximize by exclusively selecting the schedule that provides the higher return. Studies of animal choice confirm this prediction. The reason is that, given the parameters of matching in the context of consumer choice, where the schedules that govern performance are close analogs of the ratio schedules imposed in the operant laboratory, both maximization and matching theories predict a similar pattern of choice, one that eventuates in maximization and matching by virtue of the expectation that consumers will always select the cheapest alternative when selecting among brands. The expected behavior pattern is, therefore, exclusive choice of the more favorable schedule. Although there is some evidence that this is generally the case, there are frequent exceptions in that consumers sometimes buy the most expensive option or, on the same shopping trip, purchase both cheaper and dearer versions of the same product, something that animal experiments, which demand discrete choices in each time frame, do not permit their subjects. In other words, the marketing system adds complications to the analysis that cannot be anticipated within the original context of the behavioral economics research program. Even behavioral economics research with human consumers in real-time situations of purchase and consumption (token economies and field experiments) have not been able to incorporate such influences on choice as a dynamic bilateral market system of competing producers who seek mutually satisfying exchanges with consumers whose high levels of discretionary income make their selection of suppliers not only routine but also relatively cost free. Behavioral economics experiments with human consumers have at best been able to incorporate only a portion of the full marketing mix influence on consumer choice. It has typically been possible to employ price as a marketing variable but not the full panoply of product differentiation, advertising and other promotional activities, and competing distribution strategies, which are the dominant features of the modern consumer-oriented economy. Moreover, because it is the marketing mix, rather than any of its elements acting in isolation from the rest, that influences consumer choice, such experiments have been unable to capture the effect of this multiplex stimulus on purchasing and consumption.

Our analyses found that brand competition was generally marked by ideal matching, while product choices, as demonstrated here by wine and cola purchases by some degree of under-, over-, or antimatching. Relative demand curves were generally downward sloping (Foxall *et al.*, 2007). Consumers maximized by purchasing the least expensive of the brands composing their considerations sets. Where there were exceptions from the predictions of matching and maximizing theories, they occurred for reasons peculiar to the marketing context: first, because the composition of consumers' consideration sets often meant that their selections were among premium priced, higher quality brands, or at least those more highly differentiated through promotional activity, rather than among all of the brands that made up the product category.

An interesting outcome of the application of matching theory to consumer choice is the finding that brands can be defined in terms of their substitutability, while product groupings and categories whose members are independent or complementary are indicated by under-, over-, or antimatching present further evidence on this. We benefit here from making a distinction between utilitarian and informational or symbolic reinforcement, which goes beyond the usual distinction between primary and secondary reinforcers. In the case of brands, it is inevitable that they will tend to be substitutes in so far as

they are functionally similar (almost identical in terms of physical formulation), that is, in terms of utilitarian reinforcement, and complements in so far as they are differentiated by branding, that is, in terms of informational reinforcement or social symbolism. Branding is an attempt to reduce the perceived substitutability of brands by altering their value to the consumer on the basis of their social significance (e.g., increasing the status of their owners and users) or psychological significance (e.g., enhancing the self-esteem of those who own and use them).

The *pattern of reinforcement* (the pattern of low-to-high utilitarian reinforcement and low-to-high informational reinforcement produced by buying or using a product) is an analytical category that takes the place in interpretive behaviorism occupied by that of the schedule of reinforcement in the experimental analysis of behavior. Because patterns of reinforcement differ and because informational reinforcement increases the complementarity of brands within a product category, nonprice elements of the marketing mix come to the fore. The study of brand choice indicates that the multidisciplinarity of behavioral economics can usefully be extended by the inclusion of results and perspectives from marketing research. Behavioral economics is supported by the research in that its analyses and conclusions are shown to apply to human consumers in situations of free choice; behavioral economists should appreciate, however, the conclusions of marketing researchers to the effect that most consumers are multibrand purchasers, and that marketing considerations other than price influence choice. Marketing researchers may need to take note of the import of price differentials in brand choice. The behavioral mechanism of choice that underlies the molar patterns of consumer choice depicted here appears to be momentary maximization of benefit, a result that is consistent with melioration or overall maximization. However, the lesson of the research is that brand choice is reinforced by two sources of reward, *utilitarian*, which derives from the functional benefits of the good, and *informational* or *symbolic*, which derives from the psychological and cultural meanings, which goods acquire through their participation in social interactions and, by derivation, through advertising and other means to branding. The recognition of both sources of reinforcement is the key requirement for both marketing researchers and behavioral economists.

Consumer demand analysis. The point of the discussion of pattern of reinforcement and plasticity of demand is not to dismiss such stalwarts of economic analysis as elasticity of demand. Rather, having established the usefulness of the former types of interpretive construct to consumer behavior analysis, the aim should be to operationalize them and relate them to the standard constructs. This has been one of the tasks of work on consumer demand analysis within the BPM context. Much of the work on demand analysis has involved comparison of the buying patterns of consumers grouped by their predominant purchasing of brands having specific patterns of informational and utilitarian reinforcement.

Observed decreases in the quantity bought with increases in prices, indicated by negative elasticity coefficients, may be associated with different response patterns by different groups. The tendency to buy larger quantities when prices are lower may be related to one or more of the following three patterns: (i) buying larger quantities of a product when its price was below its usual, average, price rather than when its price was above its average price (i.e., intrabrand or absolute elasticity); (ii) buying larger quantities when buying brands belonging to cheaper, lower informational levels than when buying brands belonging to more expensive, higher informational levels (i.e., informational interbrand or relative elasticity); and (iii) buying larger quantities when buying brands belonging to cheaper, lower utilitarian levels than when buying brands belonging to more expensive, higher utilitarian levels (i.e., utilitarian interbrand or relative elasticity). These phenomena have been investigated in two studies, employing different sets of consumer panel data, which have borne out these extensions to matching analysis in the context of consumer choice.

The possibility of combining matching and elasticity of demand analyses has led, most recently, to the testing of an equation that

relates amount spent to quantity bought, utilitarian reinforcement obtained, informational reinforcement obtained, and price paid (which detects promotions). The results strongly support the view that economic demand is influenced not only by the amount purchased but by the variables posited by the BPM – utilitarian reinforcement, informational reinforcement, and aversive consequences – all of which can be shown to influence amount spent (Foxall *et al.*, 2007; special issue of JOBM, 2010).

Experimental analyses. Numerous studies in the experimental analysis of behavior are relevant to the development of consumer behavior analysis (see, inter alia, Foxall, 2002). However, several recent approaches to the experimental analysis of consumer behavior are of particular interest to the progress of the model and research program. The first is the use of simulated shopping malls to test predictions of matching theory in a consumer context. A series of experiments investigated the influence of various marketing elements (e.g., price and service level measured in terms of delay) on the spending behavior of buyers of consumer products. This work is significant not only for its contribution to consumer behavior analysis but also for its insights into the nature of economic psychology, that is, the manner in which the all-too-often-separated disciplines of economics and psychology, and biobehavioral consumer research might be integrated. The integration is achieved in this case through foraging theory. Although it is only one of a number of important outcomes of this kind of investigation, the finding that the price of a product may be viewed as a temporal factor by which foraging may be understood is valuable for further research that treats temporally extended consumption by humans in terms of foraging (see special issue of JoEP, 2003).

The constant need in experimental work that is designed to have some impact on the interpretation of complex behavior is to move gradually closer to empirical investigations of choice that permit field experimentation; in this way, the rigor of experimental manipulations and the links with the principles of behavior analysis can be retained while the focus of research more closely resembles the kinds of day-to-day behavior exhibited by consumers in natural settings. An interesting approach to this is found in experimental work on the influence of the base price of products on the amount of time prospective consumers spend on search. Observation of consumers during the prepurchase phase of their purchase sequence not only permits in-store methodologies to be evaluated in the context of consumer behavior analysis but also yields the result that search behavior is more extensive for higher-priced items. In-store experimentation within the consumer behavior analysis framework has also been undertaken, as well as simulated store choices to test the effects of both utilitarian reinforcement and informational reinforcement on consumer preference (JOBM, 2010).

Development

The original aim of the research program – to ascertain the epistemological status of a radical behaviorist model of consumer choice – has been fulfilled. Not only is such a model feasible, it has provided unique interpretations of consumer behavior and permitted the prediction of such aspects of consumer choice as brand and product selection. In addition, it has explained such behaviors to the extent of identifying the contingencies that shape them: both utilitarian and informational reinforcement, and the scope of the consumer behavior setting. Wider aspects of the explanation of consumer behavior have not proved amenable to this kind of theorizing, however. Such features of behavior as its continuity and its relationship to the personal level of explanation including subjective experience require the use of intentional language and this implies a different mode of explanation from the extensional approach on which radical behaviorism is based (Foxall, 2004b). The use of the BPM as an interpretive device is also enhanced by the incorporation of intentionality. This does not mean that the original formulation of the model is superseded: it presents a means of predicting and potentially influencing consumer behavior that is not available from other sources. And it fulfills the expectations of a model of consumer choice that radical behaviorism itself would entertain. But the more general demands of consumer

research as a social science can be implemented by the development of intentional and cognitive versions of the model (Foxall, 2007). The main point of these developments is to ascertain the nature of intentional and cognitive explanations of consumer choice respectively, a program that parallels the original aim with respect to behaviorist explanation. As a part of this quest, the application of the BPM to new areas such as intertemporal decision making, compulsion, impulsivity, and addiction, and the nature of the marketing firm continues apace.

Acknowledgment

The author is grateful to other key players in the research program who include Erik Arntzen, Asle Fagerstrom, Donald Hantula, Victoria James, Mike Nicholson, Jorge Oliveira-Castro, Sarah Xiao, Valdimar Sigurdsson, Teresa Schrezenmaier, and Mirella Yani-de-Soriano.

See also *choice models; consumer brand loyalty; demand elasticity; emotion; perception of brand equity*

Bibliography

Davison, M. and McCarthy, D. (1998) *The Matching Law: A Research Review*, Lawrence Erlbaum, Hillsdale.

Foxall, G.R. (2002) *Consumer Behavior Analysis: Critical Perspectives in Business and Management*, Routledge, London and New York.

Foxall, G.R. (2004a) *Consumer Psychology in Behavioral Perspective*, Beard Books, Frederick.

Foxall, G.R. (2004b) *Context and Cognition: The Interpretation of Complex Behavior*, Context Press, Reno.

Foxall, G.R. (2005) *Understanding Consumer Choice*, Palgrave Macmillan, New York.

Foxall, G.R. (2007) *Explaining Consumer Choice*, Palgrave Macmillan, New York.

Foxall, G.R. (2009) *Interpreting Consumer Choice: The Behavioral Perspective Model*, Routledge, New York.

Foxall, G.R., Oliveira-Castro, J.M., James, V.K., and Schrezenmaier, T.C. (2007) *The Behavioral Economics of Consumer Brand Choice*, Palgrave Macmillan, New York.

Herrnstein, R.J. (1997) *The Matching Law: Papers in Psychology and Economics*, Russell Sage Foundation, New York, Harvard University Press, Cambridge, MA.

Kagel, J.H., Battalio, R.C., and Green, L. (1994) *Economic Choice Theory: An Experimental Analysis of Animal Behavior*, Cambridge University Press, Cambridge.

Staddon, J.E.R. (ed.) (1980) *Limits to Action: The Allocation of Individual Behavior*, Academic Press, New York.

JOBM (2010) *Journal of Organizational Behavior Management*, **30** (2).

JoEP (2003) *Journal of Economic Psychology*, **23**, 5. Special issues of two journals contain accounts of recent theoretical and empirical advances.

family life cycle

Catherine A. Cole, Dhananjay Nayakankuppam, and Jayati Sinha

The family life cycle (FLC) emerged as a fundamental marketing concept in the 1950s. Subsequent research on the concept addressed two questions: (i) how many different stages are there in the FLC? (Wells and Gubar, 1966; Murphy and Staples, 1979; Gilly and Enis, 1982; Du and Kamakura, 2006, 2008) and (ii) how do stages in the FLC affect household expenditure, savings, and the mix of products consumed? (see Redondo-Bellon, Royo-Vella, and Aldas-Manzano, 2001 for a review).

For managers, the FLC concept has obvious practical appeal. Stages of the FLC often replace the chronological age of the head of the household as a useful segmenting variable because households within the same FLC stage spend money in similar ways, while households in different FLC stages spend money in different ways. Furthermore, because the proportion of households at different FLC changes across time in predictable ways, the FLC stage can be used as an independent variable in models forecasting primary demand for certain products such as day care or food consumption at home.

However, we suggest that there is a need to better understand the theoretical underpinnings of the FLC model. If the concept is merely an agglomeration of a set of correlated demographic variables, it has little theoretical appeal – its value mainly lies in collapsing a large number of variables into a smaller set, thereby making it more practical and useful. Thus, differences in consumption across different categories formed through the agglomeration of these demographic variables might reflect cohort effects, or shifts in consumption patterns as a function of changing circumstances (for e.g., childcare consumption, and perhaps minivan purchases, are obviously precipitated by the birth of a child). A more interesting question might be to examine shifts in consumer processes as a function of the life cycle. This is an important issue because it is not just circumstances that change and dictate consumption – there are also real social changes brought about in consumer decision processes and these could dictate consumption in much more subtle and powerful ways. More importantly, these changes are not obvious and could thus result in counterintuitive effects. For example, could the preferences of children influence the preferences of parents, and vice versa?

We start by providing a review of the extant literature on the consumption life cycle. We then review criticisms of this concept. We also propose a framework that might be useful in examining how life cycle might influence consumer processes. Finally, we report briefly on one study from our labs in which we studied how children might influence parent preferences and how household environmental variables (such as parenting style) may account for heterogeneity in preferences among households at the same stage of the FLC.

The Current Household/Family Life-Cycle Model

The assumption underlying the current FLC model is that family changes (marriage, birth of children, breakup of marriage, etc.) impact both the income and the expenditures of households. On the basis of variables such as age and labor activity of the household head, marital status, and age of youngest child, we can classify families into different stages. The stages of the FLC remain at heart a multidimensional variable resulting from combining other unidimensional ones.

Stages of the FLC. Conceptually, the FLC has evolved over time, primarily through increases in the number of stages and through increases in the flexibility of households in moving between stages. The early years established the conceptual bases of the FLC, utilizing family composition as the organizing theme. Starting in the 1940s and 1950s, the FLC models incorporated new variables as further stages were added to the model. In the more recent literature, FLC models have evolved away from the notion that households pass through an orderly progression of stages (Wells and Gubar, 1966) to the notion that households pass back and forth between stages in a more disorderly pattern (Gilly and Enis, 1982).

Table 1 summarizes the life stages of four prominent models. Wells and Gubar's (1966) model utilizes the criteria of marital status, age of

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

head of household (HH), age of youngest child, the presence of dependent children, and whether the HH is part of the labor force. The model identifies nine consecutive stages through which an individual passes: bachelor, newly married couple, full nest I–II–III, empty nest I–II, solitary survivor in labor force, and solitary survivor retired.

Murphy and Staples (1979) propose 14 stages linked by multiple paths based on marital status and presence of children. Murphy and Staples (1979) sought to incorporate nontraditional families in response to the evolving family unit. The model considered new classification variables, such as adding the category divorced in marital status and parsing some of the other variables in a more fine-grained manner (e.g., HH age was divided into three bands as opposed to the two in the Wells and Gubar model). This model allowed the age of the children to be classified as (younger than 4, from 4 to 12, and between 13 and 18) to allows substages. The modifications reduced the number of nonclassifiable households to less than 20% in the United States and less than 30% in the United Kingdom.

The Gilly and Enis (1982) model, which includes 13 stages, incorporates nontraditional unmarried cohabiting-couple households and remarriages. Like the Murphy and Staples (1979) model, the Gilly and Enis model is not a single sequence model. Instead, households can move in both directions along various routes. This model represents most HHs since it excludes only 2.8% on the same data of the US Bureau of Census. Wilkes (1995) proposed a 15-stage hybrid of the Wells and Gubar and the Gilly and Ennis typologies.

Du and Kamakura (2006) empirically developed a 13-stage model, which is summarized in Table 1. Instead of a priori defining stages, they empirically identify life stages using a hidden Markov model on household data that include characteristics such as marital status, age, and employment, and information about household size including the number of other adults and the presence of children at different ages. One interesting variable that emerges in their life-stage characterization is whether the household has children in college. Intuitively, this variable should influence household consumption because tuition often represents a major household expenditure. The data, supporting the life-stage model, were collected from a panel of approximately 8000 households in the United States that were tracked annually from 1968 through the present.

Ability of the FLC to predict consumption. Evidence is accumulating that household/FLC stage affects both the size of the consumption budget and household spending priorities for a broad variety of categories including home ownership, entertainment, energy, and other expenditures (Lansing and Morgan, 1955; Wells and Gubar, 1966; Fritzsche, 1981; Schaninger and Danko, 1993; Wilkes, 1995; Redondo-Bellon, Royo-Vella, and Aldas-Manzano, 2001; Du and Kamakura, 2008). While the reports are too numerous to review extensively, we summarize a few findings. Lansing and Morgan (1955) as well as Wells and Gubar (1966) report that income, expenditures on durable goods, assets, debts, and subjective feelings about financial position differ at different life-cycle stages. Fritzsche (1981) reports that after controlling for income and the number of people in the household, young singles in the Wells and Gubar (1966) FLC classification consume less energy than households at any other stage of the FLC (except for gasoline). Du and Kamakura (2006) analyze discretionary spending across a wide range of expenditure categories. They find that the lowest levels of expenditures are observed for older households with single, divorced, or widowed heads, while the highest are observed for large households with children. They also find allocation differences across households. For example, the top tax- and rent-paying FLC stages are households (single/divorced or young couples) with no children. In summary, there are budget and budget allocation differences across the FLC.

Criticisms of the FLC concept. While much research has verified that FLC is a discriminating variable of consumption and expenditure in a wide range of products and services, there are a number of shortcomings. There is considerable uncertainty about the generalizability of these models. They require frequent updating in the face of social trends that change the dynamics that presumably underlie the different

Table 1 Comparison of the stages of four family life-cycle models.

Wells and Gubar (1966)	*Murphy and Staples (1979)*	*Gilly and Enis (1982)*	*Du and Kamakura (2006, 2008)*
Bachelor stage; young, single, not living at home	Young (below 35) single	Bachelor I: under 35 years of age	Co/So: young single/married couple, no child, HH age 22–30
	Young (below 35) divorced, no children	—	S1: single/divorced, no child, HH age 26–42
Young, married, no children	Young (below 35) married without children	Young couple under 35 years of age	—
Full nest I, youngest child under 6	Young (below 35) married with children	Full nest I: couple, child under 6, HH under 35	C1: couple, children less than age 7, HH age 25–35
	Young (below 35) divorced with children	Single parent I: child under 6, HH under 35	S2: Divorced/single, children under age 18, HH age 27–41
Full nest II: youngest child 6+	Married, children, HH age 35–64	Full nest II: couple, children 6+, HH under 35 Single parent II: children 6+, HH under 35	C2: Family (5+), children under age 15, HH age 33–41 C4: Family (3–4), children age 7+, HH age 33–44
	Middle aged (35 to 64) divorced with children	Single parent III: HH 35–64	S3: Divorced/widow (2–3), children age 14+, college kids, HH age 44–62
	Middle aged (35 to 64) married without children	Childless couple: age 35–64	—
	Middle aged (35 to 64) divorced without children	Bachelor II: age 35–64 years	—
Full nest III: older married couples with dependent children	—	Delayed full nest: couple HH age 35–64, child under 6	C3: family (5+), children age 7+, college kids, HH age 40–50
		Full nest 3 HH age 35–64, children 6+	C5: family (3–4), children age 14+, college kids, HH age 45–57

(*continued overleaf*)

Table 1 (*Continued*).

Wells and Gubar (1966)	*Murphy and Staples (1979)*	*Gilly and Enis (1982)*	*Du and Kamakura (2006, 2008)*
Empty nest I: older married couples, no children living with them, head in labor force	Middle aged (35–64) married without dependent children	—	C6: Couple with no dependent children (age 51–73)
Empty nest II: older couple, no children at home, retired	Older (above 64) married	Older couple: 65+	C7: Family (2–3), no children under 18 or in college, HH age 63–77, retired.
Solitary survivor, working	Older (above 64) unmarried	Bachelor III: 65+	S4: Divorced/single, empty nest, HH age 49–71, working/retired
Solitary survivor, retired	—	—	S5: (Widowed) empty nest, HH age 66–84, retired

stages. Further, these models have been developed primarily within the US context and are difficult to apply in other countries with different social contexts. For example, Redondo-Bellon, Royo-Vella, and Aldas-Manzano (2001) present plausible reasons for modifying the model for the Spanish environment and develop a model with 11 stages. Wells and Gubar (1966) note that these models may not generalize across social classes because, as sociologists have noted, there may be differences in consumption goals. Finally, stages of FLC may not affect consumption of all products and services.

Other problems with the concept include the observation that no two investigators have yet agreed on the numbers and types of stages. This makes comparing the results from one study with those from another difficult. Additionally, within a single study, the researcher often has difficulty defining categories. If they are defined too narrowly, they will include a very small portion of the respondents, but if they are too broad, they will squeeze everyone into a few categories. Most models do not cover all possible types of families, nor are the identified types of families always mutually exclusive. For example, the Wells and Gubar (1966) model does not include household with children older than 17, single people older than 44, or widows with children living at home. Similarly, the Murphy and Staples (1979) model excludes families with children aged above 18, single parent households with a nondivorced parent, and those over 64 living with their children. The Gilly and Enis (1982) model treats cohabitating and married couples as the same, although their consumption patterns may differ.

Additionally, in all the models, older adults are typically lumped together in one or two stages depending on marital status. However, this categorization may mask substantial differences in the older population (those still working, living independently vs those supported by assisted living) that may account for differences in expenditures. Also, these models often do not incorporate the labor force participation of both partners.

Additionally, Commuri and Gentry (2000) raise the point that the most interesting changes in consumer behavior may be observed as households move between stages, not once they reach equilibrium in a stage. Finally, because most research on the FLC uses cross-sectional rather than longitudinal research, cohort effects may explain the observed pattern of results.

An overall problem is the focus on outcome (decisions) rather than processes (decision making). On the one hand, researchers frequently check whether the FLC model does better in predicting outcomes than socioeconomic variables such as size of the household, and characteristics of the main wage earner such as education level, job status, age, and income (Redondo-Bellon, Royo-Vella, and Aldas-Manzano, 2001). However, the models have neither identified underlying mechanisms, which explain how families make decisions, nor have they identified how these mechanisms might change across the FLC to explain different outcomes. In other words, there is little evidence suggesting that the FLC stage determines consumption.

Proposed New Model

Figure 1 contrasts the current and a proposed new FLC model. We propose a new model based on the following linkages: changed circumstances → life-cycle stage → (information processing, memory, attitudes) → behavior → consumption.

Because most people grow up in a family and continue as adults to live in social units, we believe that families have a pervasive influence on consumer behavior. We think that research in this area should shift from identifying characteristics of different life-cycle stages and consumption outcomes to analyzing how life-cycle stage affects the decision process.

Our proposed model suggests questions regarding a variety of interesting factors: how does life-cycle stage affect information processing, memory, attitudes, and behavior? For example, one difference that might emerge between different household types is the time pressure felt. It would be reasonable to predict that increases in the time pressure felt would reduce the amount of information processing for major household purchases. These changes in the time pressure felt might affect memory, attitudes, behavior, and consumption outcomes. Thus, in this example, the underlying variable of interest is not so much the stage of the FLC, but the time pressure felt by members of the household.

In the next section, we narrow down on a specific topic: the influence of college children on parents. Typical models of the FLC minimize this flow because the adult children are no longer living at home.

Example Research: College Students' Influence on Their Parents' Consumption Behavior

The study we describe represents part of a larger effort to investigate how adult children "socialize" their parents, which represents a twist on traditional socialization theory, because we are not studying how adults influence their children's brand choices but how children shape the choices of the adults. Specifically, we study the amount of attitudinal convergence between adult children and their parents for different

Current family life cycle model				
Changed circumstances in terms of partner status presence/age of children age of head of household	→	Life cycle stage	→	Consumption
Proposed new model				
Changed circumstances →	→	Life cycle stage	→	Information processing → Memory → Attitudes → Behavior → Consumption

Figure 1 The current and proposed new household/family life cycle model.

brands in different product categories. We expect that attitudinal convergence will vary across different types of product categories and across different types of families.

Product category. In our research, we investigated four different types of products: high-technology products, joint family use products, products used for individual purposes, and fast-moving consumer goods. We expect a family's opinions about brands in the high-technology products and in the joint family use categories to converge the most, but for different reasons. If a product is consumed by a family, then the family will confer about the brand and consumption is likely to be correlated – so there will likely be a convergence of attitudes. If a product is a high-technology product, then we expect the college-age student to influence the older parent – the college-age child is likely to be perceived as having expertise in the category and his/her preferences and recommendations would thus be more influential. However, for products consumed for individual purposes and for fast-moving consumer goods, it is likely that members of the household are less likely to discuss the brand choice and preferences with each other and there will thus be a divergence in family opinions. This leads to H1: attitudinal convergence will be higher for high technology products and for products that are consumed jointly and lower for frequently purchased consumer goods and items consumed for personal use. However, more important for the point we are making (namely, that these effects on attitudes are mediated through social processes), we suggest that this influence of college-age students on parents will surface only in certain kinds of households.

Family environment variables as moderators. Households within the same stage of the FLC may differ on constructs that capture family dynamics, such as nurturance, communication style, and parenting style.

Nurturance. We assessed parental warmth or nurturance using an 18-item scale with five-point Likert-type items measuring the degree to which a parent describes his/her interaction with his/her children as being warm, affectionate, and encouraging. Thus, nurturance is likely to create an atmosphere where children are more comfortable about expressing their preferences, which could thus influence parents. Thus, while we expect greater attitudinal convergence within families for technology products and joint-consumption products, we expect this to be particularly the case in families high on nurturance (H2).

Communication style. Prior research suggests that communication within the family affects children's influence. We use a communication encouragement scale, which is composed of four five-point Likert-type statements measuring the degree to which a parent believes that children should be free to candidly express their own views and disagree with their parents when they feel like it. Prior research reports that parents scoring high on the communication encouragement scale are more likely to talk to their children about consumer issues. Parents are likely to be exposed to the opinions and attitudes of their children, leading to the greater possibility of influence, depending on the extent to which children are free to express their opinions. Thus, while we expect greater attitudinal convergence within families for technology products and joint-consumption products, we expect this to be particularly the case in families high on communication style (H3).

Parental style. Parenting style refers to a group of attitudes toward the child that influence the emotional climate in which the child is raised. Prior research reports that parental style affects both children's choice of influence strategy and their degree of influence indirectly through children's perceptions of parental power. The scale can be viewed as being made up of three subscales: demandingness, emotional responsiveness, and psychological autonomy granting. The demandingness dimension reflects the extent to which parents direct their children's development through maturity, and expectations, close supervision, discipline, and confrontation when a child disobeys. The responsiveness dimension is the extent to which parents encourage their children's individuality by staying attuned to and supporting their children's needs. The autonomy-granting subscale

reflects the parental attitude toward the child having an independent point of view. Thus, while we expect greater attitudinal convergence within families for technology products and joint-consumption products, we expect this to be particularly the case in families high on parenting style, especially on the responsiveness and autonomy-grating dimensions (H4).

Method. A total of 115 student participants took part in the study for partial course credit. They proceeded to fill out a questionnaire designed to assess their attitudes toward a variety of targets as well as containing scales designed to assess family environment – parenting-style scale. Each student addressed a questionnaire to one parent. This questionnaire assessed the parent's attitudes toward the same targets as well as their assessments of the family environment – scales to measure nurturance and communication style.

Analytic approach. Since the data have a nested design (with attitudes of parents and children toward various brands nested within families), with brands at the lower level (level 1) and family characteristics at the higher level (level 2), we adopted a hierarchical modeling technique. This technique is, in essence, a mixed model incorporating fixed as well as random effects, to incorporate heterogeneity. Thus, attitudes toward any stimuli are presumed to be the result of not only characteristics at level 1 (product category) but could also result from the fact that the attitudes come from a certain social unit with interdependencies. This approach allows us to explicitly test for the hypothesized cross-level interaction, where family characteristics (at level 2) impact attitudes of both constituents (at level 1).

For the purposes of the analysis, a deviance score was calculated for each target of judgment by squaring the difference between the parent and child's attitudes. This represents the degree of convergence in family opinion – that is, the higher the deviance score, the greater the divergence in opinion (or the lower the convergence in attitudes). Thus, the deviance scores for a number of targets are nested within families that vary on the family environment variables. One can then estimate whether there is significant heterogeneity (i.e., do families differ significantly in their deviance scores across targets?). This separates the variance components across the two levels of analysis. One can then examine whether the deviance is impacted by product category, by family environment variables, and, importantly, by the cross-category interaction that has been hypothesized.

Results.

Brand attitudes. A null model yielded significant heterogeneity across families (estimate $= 5.85$, $z = 5.59$, $p < 0.0001$) and an intraclass correlation of 0.153, revealing that about 15% of the variance was at level 2 (at the level of the families). Adding the product-type variable yielded a significant main effect, F (3, 342) $= 12.64$, $p < 0.0001$. As predicted in H1, the deviance scores for technology ($M = 1.7$) and joint decisions ($M = 1.8$) were smaller than those for fast-moving consumer goods ($M = 3.7$) and personal items ($M = 2.4$). The next models added the perceptions of influence. Both parent's (F (1, 1723) $= 98.85$, $p < 0.0001$) and children's (F (1, 1724) $= 78.91$, $p < 0.0001$) perceptions of influence yielded main effects. We next proceeded to add the family environment variables.

Nurturance. Consistent with H2, analyses yielded a significant cross-level nurturance X product-type interaction, F (3, 1719) $= 3.29$, $p < 0.02$. This interaction is graphed in Figure 2, and shows that, consistent with the hypothesis, high-nurturance families have more convergent brand attitudes on the target categories than low-nurturance families. Families low in nurturance show little convergence across all categories, while families high on nurturance show a significant convergence in the target categories.

Communication encouragement style. Consistent with H3, a main effect for communication style, F (1, 113) $= 8.11$, $p < 0.01$, was qualified by a significant cross-level product-type X-communication-style interaction, F (3, 1719) $= 7.71$, $p < 0.0001$. The interaction is graphed in Figure 3 and shows that families that scored high on communication encouragement scale had higher convergence of attitudes toward brands than families that scored low on this scale in the target categories.

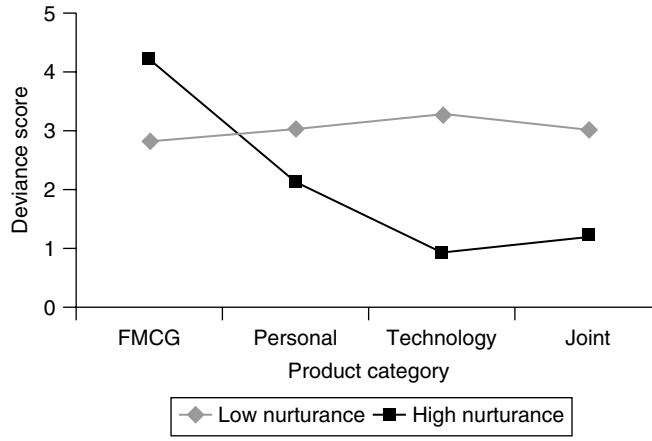


Figure 2 Nurturance × product-type interaction.

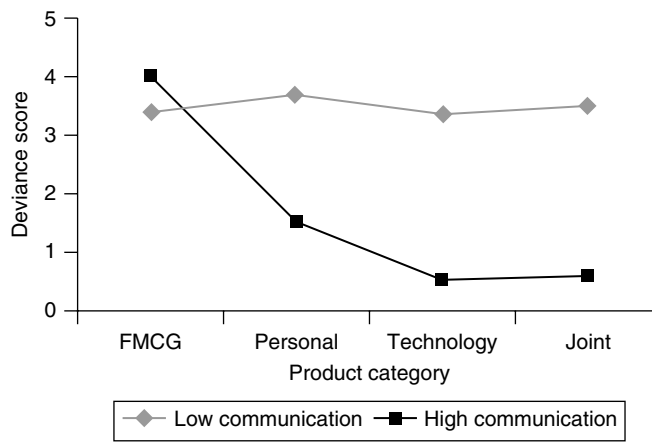


Figure 3 Communication style × product-type interaction.

Parenting style. The main effects emerged for responsiveness, ($b = -0.85$), F (1, 113) $= 3.26$, $p < 0.07$, autonomy granting, ($b = -1.3$) F (1, 113) $= 10.98$, $p < 0.0001$, and demandingness, ($b = -0.75$), F (1, 113) $= 4.89$, $p < 0.03$. As predicted in H4, these main effects were all moderated by significant cross-level interactions with product type (all interactions significant at least at $p < 0.01$ level). These interactions are graphed in Figures 4–6.

Discussion. Analyses suggest that there is considerable convergence in attitudes within families. This suggests that there are, indeed, theoretically meaningful differences in the social processes that accompany FLC stages. DINKs (double income–no kids) households, for instance, cannot be subject to the kinds of influence we have just outlined. This suggests that it might be worthwhile trying to explicate the social changes that accompany FLC stages and the impact these have on various consumer processes (information processing, memory, attitudes, etc.).

While it is difficult to parse out the direction of influence (is the convergence due to parents influencing children or children influencing parents) with the correlational data reported here, there appear to be the following reasons to think that children do influence parents' attitudes, at least to some extent. First, the

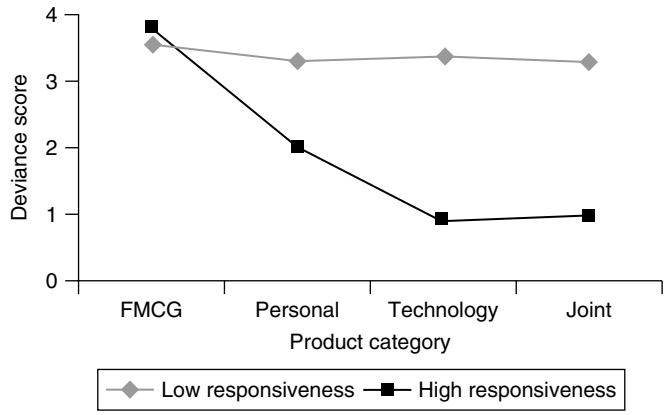


Figure 4 Parent's responsiveness × product-type interaction.

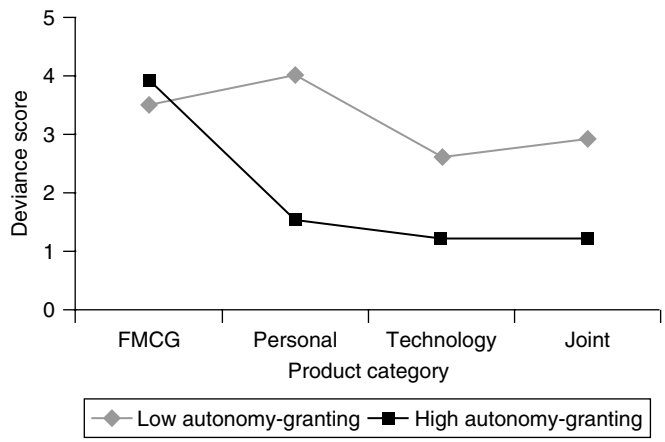


Figure 5 Parent's autonomy-granting × product-type interaction.

pattern of influence across the product types is more suggestive of children influencing parents. For example, the technology items were chosen as those that children would be more likely to have experience in (e.g., MP3 players, computers, etc.). Second, we asked participants for the extent of influence they thought they had on their parents for each of the targets of judgments. Children actually had fairly accurate perceptions of when they had an influence on their parent's attitudes – that is, their impression was that they were more likely to have influenced their parents' attitudes on the technology and joint-consumption items than the other categories. Third, the pattern of moderation by the family environment variables lends further credence to this idea – the family environment variables were chosen as theoretical constructs that would suggest a greater propensity for children to express their opinions about products – this further supports the notion that children were likely to have influenced their parent's attitudes.

Conclusion

We conclude by observing that there are several important directions for future research. First, prior research has identified meaningful parameters to use to classify households into life stages. These variables include demographics

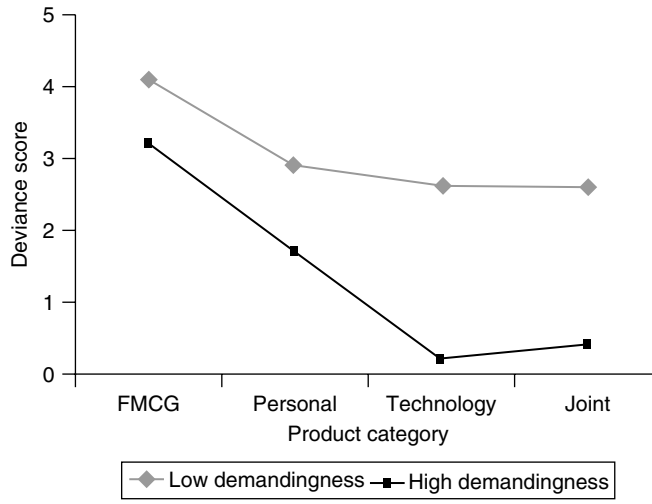


Figure 6 Parent's demandingness × product-type interaction.

such as marital status, age of the HH, employment status of the HH, and age of children (if any). Future researchers should use available sophisticated methodologies such as the hidden Markov models used by Du and Kamakura (2006) to empirically identify the most common types of households in a market and to estimate the probabilities that a household will follow an expected life path. Such an analysis could, for example, help school districts better predict enrollment trends at local elementary schools. Second, research needs to continue to identify how these emerging FLC stages influence consumption outcomes. However, such work would be enhanced if it were based on theoretically derived hypotheses about why consumption priorities shift across the FLC. Third, more research needs to be directed toward households in transition from one stage to the next stage. Households in transition may share characteristics such as a high degree of uncertainty about consumption priorities. Finally, and perhaps most importantly, more research on group consumer decision making would shed light on the processes that underlie consumption decisions made at each stage of the FLC. It is likely that such research will uncover considerable heterogeneity among families at the same life-cycle stage. Our lab study, for example, suggests that variables related to family dynamics, such as parenting style, may explain some of the heterogeneity. We believe there are many exciting unanswered questions about how family decision making changes across the FLC.

Bibliography

Commuri, S. and Gentry, J. (2000). Opportunities for family research in marketing. *Academy of Marketing Science Review*, **8**. Available at: http://www.amsreview.org/articles/commuri08-2000.pdf.

Du, R.Y. and Kamakura, W.A. (2006) Household life cycles and lifestyles in the United States. *Journal of Marketing Research*, **93**, 121–132.

Du, R.Y. and Kamakura, W. (2008) Where did all that money go? Understanding how consumers allocate their consumption budget. *Journal of Marketing*, **72**, 109–131.

Fritzsche, D. (1981) An analysis of energy consumption patterns by stage of family life cycle. *Journal of Marketing Research*, **18**, 227–232.

Gilly, M.C. and Enis, B.M. (1982) Recycling the family life cycle: a proposal for re-definition, in *Advances in Consumer Research*, vol. **9** (ed. A. Mitchell), Association for Consumer Research, Ann Arbor, MI, pp. 271–276.

Lansing, J.B. and Morgan, J.N. (1955) Consumer finances over the life cycle, in *Consumer Behavior*, vol. **2** (ed. L.H. Clark), New York University Press, New York, pp. 36–51.

Murphy, P. and Staples, W. (1979) A modernized family life cycle. *Journal of Consumer Research*, **6**, 612–638.

Redondo-Bellon, I., Royo-Vella, M. and Aldas-Manzano, J. (2001) A family life cycle model adapted to the Spanish environment. *European Journal of Marketing*, **35**, 612–638.

Schaninger, C.M. and Danko, W.D. (1993) A conceptual and empirical comparison of alternative household lifecycle models. *Journal of Consumer Research*, **19**, 580–594.

Wells, W. and Gubar, G. (1966) Life cycle concept in marketing research. *Journal of Marketing Research*, **3**, 355–363.

Wilkes, R.E. (1995) Household lifecycle stages, transitions and product expenditures. *Journal of Consumer Research*, **22**, 27–42.

consumer well-being

Aaron Ahuvia, Crystal Scott, and Elif Izberk Bilgin

Early measures of consumer well-being (CWB) equated it with a person's quantity of consumption, such that the level of well-being in a society can be summarized by its gross domestic product (GDP) per capita. In the 1960s, this reliance on GDP came under criticism by the social indicators movement, which argued for a more holistic view of human well-being beyond GDP, by including factors such as health, education, and crime rates in measures of social progress (*see* SOCIETY, CULTURE, AND GLOBAL CONSUMER CULTURE). This focus on diversifying *objective* indicators of well-being beyond GDP was further augmented in the 1970s by large-scale survey research on *subjective* indicators of well-being (SWB), such as happiness and life satisfaction. Our knowledge of subjectively experienced aspects of CWB increased dramatically in the 1990s as the rapidly growing research in positive psychology and behavioral economics brought attention to subjective well-being in all domains of life.

Income is a good proxy of a person's overall consumption level, and hence a commonly used indicator of CWB. Income, and hence consumption, is strongly linked to objective measures of well-being such as health, longevity, and education. But the link between income and subjective indicators of well-being such as happiness is much weaker and more complex (Ahuvia, 2008a, 2008b). At very low income levels where basic human needs are not met, increases in income produce lasting improvements in happiness. And at all income levels, increases in income produce *short term* increases in happiness. However, at even moderate income levels where basic needs have been met, the relationship between income and *long term sustained* happiness is extremely weak. In general, studies show that differences in income explain between 2 and 5% of individual differences in SWB, thus leaving upward of 95% of the difference between individuals in SWB to be explained by other factors such as genetics, social relationships, and the way people think about the events in their lives.

Why does increased consumption among the nonpoor fail to produce much lasting happiness? Is the problem that consumption just does not work, or is it that most people are not consuming in the right way? Spending money has been shown to produce the most lasting happiness when the money is spent on (i) charitable donations (Dunn, Aknin, and Norton, 2008), (ii) things which help foster social relationships (Lyubomirsky, 2007), and (iii) experiences as opposed to physical objects (Van Boven and Gilovich, 2003), so long as the experience was purchased with the primary goal of acquiring a life experience. Apparently, watching television is not such an experience, as it is negatively associated with CWB (Frey, Benesch, and Stutzer, 2007). Finally, there is also evidence that a good way to turn money into happiness is not to spend it at all, as savings is a good psychic investment (Headey, Muffels, and Wooden, 2008).

CWB is not only influenced by how much people spend and how they spend it, but also by their general attitudes about consumption. Higher levels of materialism are associated with lower levels of SWB (Ahuvia and Wong, 2002). Furthermore, though close social relationships are strongly linked to SWB, just thinking about money puts people in a frame of mind in which they are less inclined to reach out to others or offer others their help (Vohs, Mead, and Goode, 2006).

Marketing, and in particular advertising, has inspired controversy around its relationship to CWB (Klein, 2002; Schor and Holt, 2000). Advocates for marketing stress its role in researching consumer needs and aligning production to meet those needs (*see* MARKETING'S CORPORATE RESPONSIBILITY and STAGES OF MARKET DEVELOPMENT). These advocates also emphasize the role of marketing in generating sales, and from them profits, employment, and the other benefits of a healthy economy. Finally, they argue that advertising increases CWB by providing consumers with needed information, and by enhancing the value consumers receive from products by investing the products with symbolic meanings (e.g., coolness, masculinity, etc.) which consumers find desirable. Critics of marketing, on the other hand, argue that marketing decreases CWB by increasing

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

materialism, creating consumer desires in order to fill them, shifting consumers' priorities away from more rewarding nonconsumer activities, promoting racial and gender stereotypes, and promoting unrealistic norms for physical attractiveness or professional success which lead to disappointment and reduced SWB.

Bibliography

Ahuvia, A.C. (2008a) Wealth, consumption and happiness, in *The Cambridge Handbook of Psychology and Economic Behaviour* (ed. A. Lewis), Cambridge University Press, pp. 199–226.

Ahuvia, A.C. (2008b) If money doesn't make us happy, why do we act as if it does? *Journal of Economic Psychology*, **29** (4), 491–507.

Ahuvia, A.C. and Wong, N. (2002) Personality and values based materialism: their relationship and origins. *Journal of Consumer Psychology*, **12** (4), 389–402.

Dunn, E.W., Aknin, L.B., and Norton, M.I. (2008) Spending money on others promotes happiness. *Science*, **319** (5870), 1687–1688.

Frey, B.S., Benesch, C., and Stutzer, A. (2007) Does watching TV make us happy? *Journal of Economic Psychology*, **28** (3), 283–313.

Headey, B., Muffels, R., and Wooden, M. (2008) Money does not buy happiness: or does it? A reassessment based on the combined effects of wealth, income and consumption. *Social Indicators Research*, **87** (1), 65–82.

Klein, N. (2002) *No Logo*, Picador, New York.

Lyubomirsky, S. (2007) *The How of Happiness: A Scientific Approach to Getting the Life You Want*, Penguin, New York.

Schor, J.B. and Holt, D.B. (2000) *The Consumer Society Reader*, The New Press, New York.

Van Boven, L. and Gilovich, T. (2003) To do or to have? That is the question. *Journal of Personality and Social Psychology*, **85** (6), 1193–1202.

Vohs, K.D., Mead, N.L., and Goode, M.R. (2006) The psychological consequences of money. *Science*, **314** (5802), 1154–1156.

knowledge accessibility

Robert S. Wyer, Jr.

General Considerations

Definition. Knowledge is accessible in memory if it comes to mind quickly and easily in the course of pursuing a goal to which it is relevant. As such, it is distinguished from *availability*, which refers to the existence of knowledge in memory. The distinction is important in light of the fact that people are normally neither motivated nor able to bring all of the knowledge they have accumulated that is relevant to a judgment or decision they make. Rather, they retrieve and use only a subset of this knowledge that comes easily to mind, ignoring other knowledge that is equally applicable but less easily accessible.

Theoretical formulations of the determinants and consequences of knowledge accessibility depend on the type of knowledge in question. *Declarative* knowledge includes single concepts that are used to interpret individual pieces of information, configurations of attributes that are used in combination as a basis for identifying a stimulus as a member of a more general category, beliefs, and attitudes. It can also include sequences of temporally and thematically related events. Some sequences characterize specific experiences one has had and constitute stories that are retrieved and communicated to others. Other, more general (prototypic) sequences can be used to infer the cause of a new event or to predict its likely consequences. These latter sequences can constitute *implicit theories* about the social or physical world. Other sequences may specify the steps required to attain a particular objective.

In addition, *procedural* knowledge concerns the sequences of cognitive or motor behavior that have been acquired through learning and are performed automatically when the configuration of stimuli with which they are associated are experienced. These procedures can include those that operate on declarative knowledge in the course of attaining a particular objective.

Determinants of accessibility. The accessibility of knowledge in memory typically increases with both the recency and the frequency with which it has been activated and used in the past. Individuals who search memory for knowledge that is relevant to a goal they are pursuing presumably compile a set of features that specifies the type of knowledge required, and these features cue the retrieval of previously acquired units of knowledge that have the features in common.

If concepts and knowledge have been used recently, they typically come to mind more quickly and easily than other, equally relevant concepts and knowledge. Therefore, they are more likely to be used in attaining a goal to which they are relevant. For example, a person may evaluate a luxury car more favorably if (s)he has recently been asked to judge the aesthetic appeal of a series of paintings (thereby increasing the accessibility of aesthetics-related criteria) than if (s)he has recently been asked assess the speed and memory capacity of a computer (thus activating concepts associated with functional utility).

However, the likelihood of knowledge coming to mind in the course of goal-directed activity is also a function of the *frequency* with which it has been used in the past. Thus, concepts and knowledge that have been used very frequently over a period of time become *chronically* accessible in memory and, in the absence of other considerations, may be applied rather than other, potentially applicable but less accessible cognition.

Underlying processes. The situational and individual difference factors that determine the accessibility and use of knowledge are well documented. However, ambiguities continue to surround the cognitive processes that underlie their effects. These processes depend on both the type of knowledge in question and the theory of memory one assumes.

The processes that underlie the accessibility of declarative knowledge are most commonly conceptualized with reference to a spreading activation model of associative memory (Collins and Loftus, 1975; but see Wyer and Srull, 1989, for an alternative). According to this model, concepts and units of knowledge are represented in memory by nodes and the associations between them by pathways that connect them. When a unit of knowledge is activated (that is, thought

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

about), excitation spreads from the node at which it is located to other nodes along the pathways connecting them. When the excitation that has accumulated at one of these nodes exceeds a certain threshold value, the knowledge at this node is activated (comes to mind). When this knowledge is no longer thought about, the excitation that exists at the node dissipates. However, it takes time to dissipate entirely. Therefore, if a relatively short period of time has elapsed, residual activation may still exist at the node. Consequently, the knowledge is more likely to be reactivated by excitation that is transmitted to it from other sources (e.g., new stimulus information, or a directed search of memory for goal-relevant knowledge).

These processes govern the conscious use of declarative knowledge in the course of goal-directed activity. They also capture the conscious identification and use of a goal-directed procedure under conditions in which more than one strategy is potentially applicable. However, cognitive and motor procedures are often applied automatically, with little, if any, deliberation. These procedures, which constitute procedural knowledge, can be conceptualized as a set of *productions*, or learned "if [X], then [Y]" rules, where [X] is a configuration of externally or internally generated stimulus features and [Y] is a sequence of cognitive or motor acts that are elicited automatically when the preconditions specified in [X] are met (Anderson, 1983).

Productions can guide the cognitive and motor activity in a large number of situations, ranging from the generation of speech to driving a car or using a word processor. However, the activation of a production depends on the particular set of concepts and declarative knowledge that is accessible in memory at the time and, perhaps fortuitously, happens to be included among the features that compose a production's precondition [X]. Thus, to borrow an example from Bargh, Chen, and Burrows (1996), experimental participants who fortuitously have the concept "slow," accessible in memory as a result of thinking about the elderly, are likely to walk more slowly to the elevator upon leaving the experiment.

Methodological Considerations

The effect of knowledge accessibility on judgments and behavior has generally been investigated using *priming* techniques, that is, participants are typically asked to perform a task or engage in activity that requires the use of a particular concept, thus increasing its accessibility in memory. The effects of this activity on judgments and behavior in an unrelated situation are then observed. Furthermore, concepts and knowledge can often be activated subliminally in the course of performing an unrelated task.

However, concepts and knowledge can also become chronically accessible in memory as a result of their frequent application over an extended period of time. Thus, for example, cultural differences in the interpretation of information and the construal of its implications can be a reflection of differences in the chronic accessibility of concepts and knowledge that result from frequent exposure to them in the course of growing up. Although the effects of priming concepts and knowledge on their accessibility in memory can temporarily override the effects of their chronic accessibility, the latter effects are likely to predominate after a period of time has elapsed.

The role of awareness. Individuals are not always conscious of why a piece of knowledge comes to mind, and are likely to attribute it to the type of object they are judging or the decision they are called upon to make. Consequently, when two or more units of knowledge are equally useful in attaining a particular objective, objectively irrelevant factors that influence their relative accessibility in memory can determine which is applied, and this can occur without awareness that the influence has occurred.

In fact, when individuals *are* aware that concepts and knowledge come to mind for reasons that have nothing to do with the stimuli being judged, they may intentionally avoid using them. As a result, the cognitions are less likely to be applied than they might otherwise be. On the other hand, conscious attempts to suppress a concept or unit of knowledge require thinking about the knowledge to be suppressed.

Thus, ironically, knowledge may become more accessible in memory as a result of the attempts to suppress it, and, consequently, it may be *more* likely to be used once active attempts to suppress it no longer exist. Thus, for example, individuals who actively avoid using a person's ethnicity or social status as a basis for judgment in order to comply with sanctions imposed in one situation are more likely to use the stereotype in responding to a different person once the sanctions are lifted (Bodenhausen and Macrae, 1998).

Individuals are normally aware that they are applying certain criteria in arriving at a judgment or decision, but are nonetheless unaware of why they applied these particular criteria rather than others that are equally applicable. However, behavior of the sort that is governed by productions can be activated and applied without awareness of the behavior itself. Thus, as in the earlier example, activating concepts associated with the elderly may stimulate people to walk slowly without awareness of their walking speed. Similarly, if a driver who is on the way to the store happens to think about a meeting she has the next day, these thoughts may activate a product that leads her to wind up in front of her office rather than at her intended destination.

Applications

The effects of stimulus information on behavior and decisions can be localized at several stages of processing, including the attention to information, its comprehension, inference, and evaluation, and the generation of an overt response or behavioral decision. The following review of representative research on the effects of knowledge accessibility at each stage of processing provides an indication of the range of phenomena in which knowledge accessibility plays a role. (For a more detailed discussion of this research, see Förster and Liberman, 2007; Wyer, 2008.)

Attention and search processes. Effects of declarative knowledge accessibility. When individuals receive specific items of information about an object for the purpose of making a judgment, they may first interpret information in terms of more general concepts that the information exemplifies. However, information that can be interpreted in terms of concepts that are accessible in memory are likely to be more quickly identified and encoded in terms of them than is information that is less easily interpreted. This selective encoding can influence the judgments that are ultimately made.

The concepts that are accessible in memory are likely to be influenced by the goal that individuals are pursuing at the time. The selective attention to information that exemplifies these concepts can influence judgments of the object to which the information refers. For example, individuals who receive information about a person's behavior for the purpose of deciding whether the person is an extrovert may activate attributes of an extrovert and identify information that exemplifies these attributes, ignoring information that describes attributes of an introvert. Consequently, they are likely to judge the person as more extroverted than might otherwise be the case. Similarly, consumers who are asked to decide which of several products they like most may focus their attention on favorable attributes, whereas those who are asked to decide which they dislike most are more likely to focus on unfavorable attributes. Thus, the first group of subjects is likely to judge the products as a whole more favorably than the second group.

Attributes that happen to be accessible in memory for goal-irrelevant reasons can have similar effects. For example, exposing individuals to news stories that concern either domestic or foreign affairs can affect the criteria that individuals used to assess the performance of the United States President. Similarly, priming concepts associated with either taste or health can influence the criteria that consumers use to evaluate a product they encounter in an ostensibly unrelated situation.

Effects of activating procedural knowledge. Other studies show that the procedures that individuals use to search for information can be influenced by the procedures they have employed in the course of performing other, unrelated activities. For example, individuals who have previously rank ordered stimuli from high to low along a dimension, which requires a consideration of stimuli with high values before stimuli with low values, use a similar strategy

when scanning an array of information they encounter later (i.e., the prices of hotels in a given city, or individual consumers' ratings of a product's quality). Consequently, they make higher estimates of the implications of the information than do individuals who have previously ranked stimuli from low to high (Shen and Wyer, 2008). (This, of course, occurs only when individuals do not have the time or motivation to consider all of the information available.)

Comprehension processes. Effects of declarative knowledge accessibility. When people receive information that can be interpreted in different ways, the interpretation they give to it is likely to be determined by the applicable concepts that are most accessible in memory at the time. Thus, for example, exposing individuals to the words such as "adventurous" or "reckless" in the course of performing an ostensibly unrelated color-naming task can influence the concept they use to interpret information that a person wants to cross the Atlantic in a sailboat and, therefore, can influence the evaluation of this person. Similar effects can occur when the concepts used to interpret the behavioral information are primed subliminally.

Accessible concepts and knowledge can have more general effects. For example, statements that appear anomalous when encountered out of context (e.g., "The haystack was important because the cloth would rip") are given meaning, and consequently are better remembered, by preceding them with a single word (e.g., "parachute"). The word apparently activates a complex body of knowledge that permits recipients to construct a mental picture of the situation in which the statement makes sense.

A consideration of the impact of knowledge accessibility on the interpretation of information increases in importance by virtue of the fact that once this interpretation is made, it is used as a basis for later judgments of the information's referent independently of the information itself. Moreover, this tendency increases over time. For example, priming a concept of hostility may lead a target person to be judged as hostile on the basis of behavior that exemplifies this concept. However, this effect is greater a day after the original information is presented than it is immediately afterwards. Similarly, once a person has been judged as honest on the basis of information that he told his girlfriend that her hairstyle was ugly, the person is judged as more kind than he would be if this initial judgment had not been made, and this tendency also increases over time (Carlston, 1980).

People's reliance on their interpretation of information rather than the information itself can produce distortions of memory. Research on eyewitness testimony (Loftus, 1975) indicates that individuals who have been shown a picture of an automobile accident make higher estimates of the car's speed if the question they are asked refers to the car "smashing into the tree" than if it refers to "hitting the tree." Furthermore, they are more likely to mistakenly remember that there was broken glass in the picture in the first case. Participants apparently reconstruct a mental image of the accident that is consistent with implications of the verbal description they are given and later use this reconstruction as a basis for their recall rather than the original picture they saw.

The chronic accessibility of concepts and knowledge can also affect the interpretation of information. The accessibility of these concepts can arise in part from the frequency of using them in the course of daily life. Thus, for example, music majors are more likely than physical education majors to interpret an ambiguous passage as pertaining to the rehearsal of a string ensemble rather than a game of cards. However, the physical education majors are relatively more likely to interpret a passage as being about a wrestling match rather than a jail break (Anderson *et al*., 1977).

Assimilation and contrast effects. Primed concepts are more likely to be used as a basis for interpreting information only when the information is sufficiently ambiguous that the concept can plausibly be applied to it. If the possible meanings the information can have do not fall within the range to which the concept is applicable, the concept may be used as a standard of comparison, leading it to have a contrast effect. Thus, for example, exposing participants to the names of moderately hostile individuals may lead a target's ambiguous behavior to be interpreted as more hostile than it would otherwise be, but exposing them to

extreme exemplars of hostility (e.g., Adolf Hitler) may lead the behavior to be interpreted as *less* hostile than it would otherwise be. Correspondingly, consumers typically judge the price of a product, which is unambiguous, to be less expensive if they have previously been exposed to high prices than if they have been exposed to low prices. Furthermore, this is true even when the primed values are conveyed subliminally (Adaval and Monroe, 2002).

Effects of procedural knowledge accessibility. People may employ different strategies in comprehending information. Some individuals, for example, may have a disposition to interpret information in terms of semantic concepts regardless of the modality in which the information is conveyed. Others, however, may be disposed to construct visual images on the basis of the information even when the information is conveyed verbally. These dispositions, which can be either chronic or situationally primed, can influence the impact of the information. For example, individuals with a disposition to form visual images from verbal information may find it difficult to construct an image of an unfamiliar product and to organize descriptions of its attributes into this image. Thus, they may evaluate the product less favorably persons who process the information semantically.

Furthermore, individuals with a disposition to form visual images may have difficulty processing ad information about an object (e.g., a hotel resort) if it is described from different perspectives (e.g., that of someone inside vs outside the hotel) and may evaluate the object less favorably than those who interpret the individual items in terms of semantic concepts without forming visual images.

A second processing style with implications for consumer judgment and decisions concerns the disposition to comprehend information items individually or in relation to one another or the context in which they are found. This disposition, like the disposition to process information verbally or visually, can be either chronic or situationally induced. For example, stimulating people to think of themselves as individuals or in relation to others can be induced by asking them to use either "I" or "we" repeatedly in a sentence construction task. Once activated, these self-referent processing styles can influence the processing of information in quite unrelated domains. For example, inducing individuals to think of themselves in relation to others increases their memory for the positions of physical objects in a stimulus array independently of memory for the items themselves. It also increases their memory for contextual features of a stimulus array as well as the focal object.

The effects of priming independent versus relational thinking are paralleled by the effects of chronic differences in the accessibility of these thinking styles. North Americans typically think of themselves as independent of others, whereas Asians tend to think of themselves in relation to others. These chronic differences in self-construal, like situationally induced differences, are manifested in more general differences in thinking style (Nisbett, 2003). Thus, when asked to group stimuli (e.g., a man, a woman, and a child), Westerners are likely to group them on the basis of category membership (putting the man and woman together, as both are adults) whereas Asians are more likely to group them on the basis of their relation to one another (e.g., to group the woman with the child because the woman takes care of the child). Furthermore, although Asians and Westerners are equally able to identify and remember differences in the central features of a picture, Asians are more likely to remember peripheral, contextual features of the stimuli as well.

Inference and evaluation. Beliefs. When individuals are asked to report their belief that a statement is true, they are likely to search memory for previously acquired knowledge that has implications for its validity. However, they normally retrieve and use only a small subset of their knowledge that comes to mind quickly and easily. This knowledge may often include a semantically equivalent statement they have encountered at an earlier point in time. However, although individuals often remember having encountered such a statement before, they are less likely to remember where or when they did so. In such instances, they may base their judgment on the statement's familiarity. This, in turn, may depend on the ease with which the statement comes to mind.

Thus, if knowledge bearing on a statement's validity has been encountered recently and, therefore, is easily accessible in memory, individuals may often infer that the statement is true without considering the source of this information. For example, people are likely to judge a proposition to be true if they encountered it in a questionnaire they had completed a few days before. Similarly, they are likely to judge a fictitious person to be well known if his/her name was encountered in a different context 24 hours earlier. Perhaps more interesting is evidence that when consumers have been repeatedly exposed to an advertising claim along with an assertion that the claim is false, they later report this as more likely to be true than they would if they had never been exposed to it at all.

When a previously acquired representation of the statement to be judged does not come quickly to mind, individuals are likely to search for other relevant knowledge that bears on its validity. In doing so, they may retrieve and use only a subset of this knowledge that is most accessible in memory, assuming that it is representative of the larger subset that they have acquired and not considering other information that may be applicable. Thus, to give a simple example, they may report a stronger belief that drinking coffee is desirable if they are asked in the morning (when the knowledge that it wakes you up is easily accessible) than if they are asked late at night (when the fact that it gives you insomnia is more accessible).

Individuals' estimates of the likelihood of an event can often be based on their perception of the frequency of its occurrence and this perception, in turn, may depend on the ease with which exemplars of the event come to mind. Shrum, Wyer, and O'Guinn (1998) provide evidence of this in research on the effects of watching television on perceptions of social reality, that is, individuals typically overestimate the incidence of objects and events in the real world when the events occur frequently on television. Furthermore, this tendency increases with the amount of television they watch. Thus, for example, heavy television viewers are more likely than light viewers to overestimate the incidence of violent crime, the number of lawyers and doctors, and the number of individuals who have swimming pools in their backyard.

The assumption that these effects are mediated by the ease with which instances come to mind is strengthened by evidence that when instances are difficult to generate, they can have a negative effect on judgments that is independent of the number that are actually identified (Schwarz, 2004). Thus, for example, individuals report less favorable evaluations of a product after generating many favorable attributes of it (which is difficult to do) than after generating only a few such attributes (which is easy).

Attitudes. Attitudes toward an object, event, or state of affairs are often inferred from estimates of its desirability. Like beliefs, they can sometimes be based on the subset of knowledge that comes to mind most quickly. Although this knowledge can often include a previously formed evaluation of the stimulus in question, other judgment-relevant knowledge can be retrieved and used as well. To this extent, the stability of an attitude depends on whether the same or a different subset of previously acquired knowledge happens to come to mind each time the attitude is reported. In some cases, a recent behavior that has implications for one's attitude may be retrieved and used independently of other criteria. Thus, individuals who have recently volunteered to advocate a position on an issue may later recall this behavior and infer that they favor the position without considering other attitude-relevant knowledge they have acquired (Bem, 1972).

A major source of information about one's attitude toward a stimulus can be one's affective reactions to it. That is, an individual can have positive or negative reactions to a stimulus as a result of encounters with it in the past, and thinking about the stimulus can reelicit these feelings. These feelings can then be used as a basis for evaluating it. However, individuals are often unable to distinguish clearly between the affect that is actually elicited by an object they are judging and the feelings they are experiencing for other, perhaps irrelevant reasons. Consequently, extraneous affect that is elicited by thinking about an unrelated past experience, or by objectively irrelevant features of the situation in which an object is judged (e.g., the weather, or music that happens to be playing) can influence judgments of the object (Schwarz and Clore, 1996). In the product domain, the

affect that is elicited by a picture of a product can influence consumers' evaluations of the product independently of more specific attribute information that they encounter later (Yeung and Wyer, 2004).

Decision processes. The implications of the effects of knowledge accessibility for consumer behavior are particularly evident at the decision stage of processing. Research on these effects has concerned both the criteria that individuals bring to bear on their decisions and the procedures they employ when making these decisions.

Effects of declarative knowledge. When the products that consumers consider purchasing have both desirable and undesirable attributes, their choices can depend on which type of attribute they weight more heavily. Suppose one product has a very favorable attribute but also a very unfavorable one. However, a second has a moderately favorable attribute and a moderately unfavorable one. In this case, consumers who focus on favorable attributes are likely to choose the first product, whereas consumers who are concerned with negative attributes are likely to choose the second.

These dispositions may be either chronic or induced by situational factors that are quite irrelevant to the decision to be made. For example, calling individuals' attention to a discrepancy between their actual self concept and their ideal self appears to stimulate a *promotion* focus, that is, a tendency to think about positive consequences of their behavior. In contrast, calling individuals' attention to a discrepancy between their actual self and standards established by others stimulates them to think about avoiding negative consequences. The latter, *prevention* focus can also be induced by increasing individuals' awareness of their membership in a group, thus activating thoughts about social responsibility and decreasing the willingness to take risks. Once activated, these dispositions influence choice behavior in situations that are unrelated to the situations that gave rise to them.

Cultural differences in the tendency to focus on positive or negative decision consequences have also been identified, that is, North Americans are disposed to focus on the positive features of choice alternatives, whereas Asians are more concerned about the avoidance of negative features. However, these cultural differences may only be apparent if situational factors (e.g., the language in which individuals are communicating, or the need to give a reason for their choices) increase the accessibility of culture-related norms and values (Briley, Morris, and Simonson, 2000, 2005).

Effects of procedural knowledge. Making a purchase decision can involve three steps: deciding whether to make a purchase, deciding which of several alternatives to buy, and deciding how to implement the purchase (e.g., to pay by cash or credit card). Although these steps often occur in sequence, increasing the accessibility of later steps in the sequence can sometimes lead consumers to apply these steps without performing the earlier ones. For example, consumers may think about how to pay for a product without thinking about whether they actually want to purchase it. Similarly, they may consider which product to choose without thinking about whether they want to buy anything at all. To this extent, activating concepts associated with later stages in the decision process may lead consumers to apply these concepts without performing earlier steps, and this can affect their likelihood of making a purchase.

Thus, for example, consumers who are induced to make a small purchase early in an experimental session are more inclined to make a second purchase later in the session than they otherwise would be. Furthermore, inducing consumers to state a preference for one of two choice alternatives can induce a "which to choose" mind-set that increases their likelihood of making a purchase without considering the option of buying nothing at all. Moreover, this is true even if the preferences they report initially pertain to products that are quite different from the ones considered in their later purchase decision and, in fact, do not have to concern products at all (that is, reporting preferences for wild animals, like reporting preferences for products, can increase the likelihood of purchasing snacks that are on sale after the experiment; see Xu and Wyer, 2008).

Overt behavior. The subset of knowledge that is accessible in memory can influence overt behavior as well as judgments and decisions. Moreover, this influence can occur without awareness. A number of studies by Bargh (1997) and his colleagues provide evidence. For example, individuals who have been primed with concepts associated with rudeness in a sentence construction task are more likely to interrupt an experimenter's conversation with a graduate student in order to return a questionnaire. Similarly, individuals who have been primed with concepts associated with the elderly walk more slowly to the elevator after leaving the experiment. Subliminally exposing Caucasian students to faces of African Americans (who are stereotypically associated with aggressiveness) increases their nonverbal manifestations of irritation when they were asked to repeat a boring task. These and other studies suggest that priming semantic concepts activate a production that elicits overt behavior without awareness.

Bibliography

Anderson, J.R. (1983) *The Architecture of Cognition*, Harvard University Press, Cambridge.

Anderson, R.C., Reynolds, R.E., Schallert, D.L., and Goetz, E.T. (1977) Frameworks for comprehending discourse. *American Educational Research Journal*, **14**, 367–381.

Adaval, R. and Monroe, K.B. (2002) Automatic construction and use of contextual information for product and price evaluations. *Journal of Consumer Research*, **28**, 572–588.

Bargh, J.A. (1997) The automaticity of everyday life, in *Advances in Social Cognition*, Vol. 10 (ed. R.S. Wyer), Erlbaum, Mahwah, pp. 1–62.

Bargh, J.A., Chen, M., and Burrows, L. (1996) Automaticity of social behavior: direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, **71**, 230–244.

Bem, D.J. (1972) Self-perception theory, in *Advances in Experimental Social Psychology*, Vol. 6 (ed. L. Berkowitz), Academic Press, New York, pp. 1–62.

Bodenhausen, G.V. and Macrae, C.N. (1998) Stereotype activation and inhibition, in *Advances in Social Cognition*, Vol. 11 (ed. R.S. Wyer), Erlbaum, Mahwah, pp. 1–52.

Briley, D.A., Morris, M., and Simonson, I. (2000) Reasons as carriers of culture: dynamic versus dispositional models of cultural influence on decision making. *Journal of Consumer Research*, **27**, 157–178.

Briley, D.A., Morris, M.W., and Simonson, I. (2005) Cultural chameleons: biculturals, conformity motives and decision making. *Journal of Consumer Psychology*, **15**, 351–362.

Carlston, D.E. (1980) Events, inferences and impression formation, in *Person Memory: The Cognitive Basis of Social Perception* (eds R. Hastie, T. Ostrom, E. Ebbesen *et al.*), Erlbaum, Hillsdale, NJ, pp. 89–119.

Collins, A.M. and Loftus, E.F. (1975) A spreading-activation theory of semantic processing. *Psychological Review*, **82**, 407–428.

Dijksterhuis, A. and Bargh, J.A. (2001) The perception–behavior expressway: automatic effects of social perception on social behavior, in *Advances in Experimental Social Psychology*, Vol. 33 (ed. M.P. Zanna), Academic Press, San Diego, pp. 1–40.

Förster, J. and Liberman, N. (2007) Knowledge activation, in *Social Psychology: Handbook of Basic Principles*, 2nd edn (eds A.W. Kruglanski and E.T. Higgins), Guilford, New York, pp. 201–231.

Higgins, E.T. (1996) Knowledge activation: accessibility, applicability, and salience, in *Social Psychology: Handbook of Basic Principles* (eds E.T. Higgins and A. Kruglanski), Guilford, New York, pp. 133–168.

Loftus, E.F. (1975) Leading questions and the eyewitness report. *Cognitive Psychology*, **7**, 560–572.

Nisbett, R.E. (2003) *The Geography of Thought: How Asians and Westerners Think Differently*, Free Press, New York.

Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology*, **14**, 332–348.

Schwarz, N. and Clore, G.L. (1996) Feelings and phenomenal experiences, in *Social Psychology: A Handbook of Basic Principles* (eds E.T. Higgins and A. Kruglanski), Guilford, New York, pp. 433–465.

Shen, H. and Wyer, R.S. (2008) Procedural priming and consumer judgments: Effects on the impact of positively and negatively valenced information. *Journal of Consumer Research*, **34**, 727–737.

Shrum, L.J., Wyer, R.S., and O'Guinn, T. (1998) The effects of watching television on perceptions of social reality. *Journal of Consumer Research*, **24**, 447–458.

Taylor, S.E. and Fiske, S.T. (1978) Salience, attention and attribution: top of the head phenomena, in *Advances in Experimental Social Psychology*, Vol. 11 (ed. L. Berkowitz), Academic Press, New York, pp. 249–288.

Wyer, R.S. (2008) The role of knowledge accessibility in cognition and behavior: implications for consumer information processing, in *Handbook of Consumer Psychology* (eds C. Haugtvedt, P. Herr, and F. Kardes), Erlbaum, Mahwah, pp. 31–76.

Wyer, R.S. and Srull, T.K. (1989) *Memory and Cognition in its Social Context*, Erlbaum, Hillsdale, NJ.

Xu, A.J. and Wyer, R.S. (2008) The comparative mindset: from animal comparisons to increased purchase intentions. *Psychological Science*, **19**, 859–864.

Yeung, C.W.M. and Wyer, R.S. (2004) Affect, appraisal and consumer judgment. *Journal of Consumer Research*, **31**, 412–424.

possessions and self

Russell Belk

The concept of extended self (Belk, 1988) posits that there is first an atomized individual self that radiates out into the world by means of both tangible possessions and other people to whom one feels connected. This notion can be traced to William James (1890) who observed that

> A man's Self is the sum total of all that he CAN call his, not only his body and psychic powers, but his clothes and his house, his wife and his children, his ancestors and friends, his reputation and works, his lands, and yacht, and bank account. All these things give him the same emotions. If they wax and prosper, he feels triumphant; if they whither and die away, he feels cast down, – not necessarily in the same degree for each thing, but in much the same way for all (291–292).

Belk (1987, 1988) presents various primary and secondary evidence demonstrating that we regard our possessions as constitutive parts of our selves and that both things and people are commonly perceived as varying in their centrality as components of our self conception. When our key definitional objects are people, we construe the self in a more aggregate way. Thus, our children's accomplishments or setbacks are felt as our own and our parents' behaviors can be a source of pride or shame for us as children. Aggregate extended self provides one mechanism to explain why we are more willing to share with those whom we regard as a part of us ("sharing in") than with others ("sharing out") (Belk, 2010). At a more extended aggregate level it helps to understand why our sports team's, university's, or nation's achievements or failures can bring us pleasure or pain.

A related concept is that of attachment to possessions (Belk, 1992; Kleine and Baker, 2004). Because certain possessions are seen as more central to our sense of self, they are protected more, shared less, cared for better, and mourned more when they are lost or damaged. Even internal bodily organs are less likely to be donated when they are seen as more central to our identity (Belk, 1987). An additional individual characteristic that is related to unwillingness to share possessions is materialism (Belk, 1985; CONSUMER MATERIALISM). *Materialism* is the belief that having or not having desired possessions is a primary source of happiness or unhappiness in life. Theorized subcomponents of materialism include nongenerosity and possessiveness, both of which help explain why more materialistic people may be less generous in sharing their possessions with others. Ironically, despite materialists' beliefs that possessions contribute to feelings of well-being, those who are more materialistic tend to be less happy and report lower feelings of well-being than those who are less materialistic. Although it might seem that those who eschew acquiring or upgrading possessions in favor of experiences might be less attached to possessions, experiences too can be regarded possessively and as self-extending. This is the case, for instance, when we claim bragging rights to places we have visited and things we have done. And even those who practice voluntary simplicity may be attached to those things they own, perhaps even more than those who have more things, but regard them more superficially. In this sense, a wasteful society of abundance may cause a problem in which we are insufficiently materialistic rather than overly so.

Possessions and Identity

It should not be assumed however that possessions reflect on us only in terms of their quantity or monetary worth. Rather, the choice of particular possessions and groups of possessions help to define our lifestyles, group affiliations, and perceived personalities. Because the meanings of possessions are defined socially as well as by marketing in an age of branded and increasingly global commodities, the role of particular possessions in identity and impression formation is a changing one. Fashions, novelty, affordability, and the adoption patterns of others all influence these image-related aspects of possessions. Moreover, it is seldom a single good that provides a recognizable identity, but instead a particular ensemble or constellation of goods allows recognition of coherent patterns that can be interpreted as conveying consumption meanings. It takes a coherent and recognizable set of possessions to make an intelligible statement about our identity. In addition, it is not only

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

ownership of certain things or affiliations with certain people that conveys meanings. If the consumer does not have the requisite knowledge or skill to deploy these goods or interact with these people effectively, their possession or affiliation will fail to allow us to bask in their glory. For example, knowing the right wines to order or having the right acquaintances may not be enough to convey sophistication if the consumer is unable to master a meaningful conversation about viniculture or topics of interest to the people who might otherwise enhance our image by virtue of our association with them.

One important study of favorite self-expressive possessions was a three-generation study carried out in Chicago by Csikszentmihalyi and Rochberg-Halton (1981). The youngest, largely teenaged, generation treasured things that helped them perform (e.g., musical instruments, sports equipment). Their 40- to 50-year-old parents favored things that showed their status or power (e.g., automobiles, gadgets). And their grandparents cited objects that showed their connections to friends and family members (e.g., photographs, gifts from others). While this was a cross-sectional study, other studies have found similar life-cycle differences in the meanings of things as well as the degree of materialism exhibited by different generations (*see* FAMILY LIFE CYCLE). Earlier in the life cycle, possessions may help the infant to define a separate identity from his or her mother and subsequently possessions may act as transitional objects like a security blanket to ease the separation from mother and home. Later in the life cycle, as we prepare for death, there is often an attempt among the middle class of the world to leave a legacy for children via heirlooms and other mementos that hopefully leave a trace of our life behind for subsequent generations (e.g., Price, Arnould, and Curasi 2000).

Because of their functions in potentially creating meaningful identities and links to other people over the life cycle, possessions in a society without fixed and inherited roles also help us to envision and enact different selves. Some of these different selves are situational, as when we don different shoes, clothes, and cosmetics to go to work, go to a play, or go running. But there are other self transformations and possession-related rites of passage that ease us through transitions related to aging, educational and career milestones, marriage, divorce, parenthood, sickness, mourning, and other changes in status, health, appearance, employment, and social connections. In cultures that value independence and self-determination, our desire for expressing uniqueness may be stronger than our desire for expressing affiliation. But normally, both of these forces shape our behavior and both are manifest in the products, brands, and experiences that we acquire as consumers. In efforts to defy aging or try on new personas we may change wardrobes, visit beauty spas, acquire cosmetic surgery, accumulate travel experiences, or simply change brands.

Because possessions are often central to our identity, the involuntary loss of possessions to theft or disaster is often traumatic. Wicklund and Gollwitzer (1992) theorize that at such times as well as when we otherwise feel a lessened sense of self, we attempt to shore up or reconstitute our identity through acquiring new possessions as an act of symbolic self completion. They found, for example, that MBA students who were less secure about their job prospects were more likely to attempt to compensate by acquiring stereotypical business artifacts like new shoes and suits, expensive pens, and attaché cases. This can be seen as a case of trying to improve the image conveyed by the extended self and what we have rather than by what we do or the person we are in some deeper sense.

One approach to understanding the role of possessions in identity is to ask people to describe themselves. Although there are cultural differences in the responses people give to this question, in the West it might be common to list our name, age, sex, occupation, religion, and hometown first. But after that we might begin to list our consumption interests (e.g., musical tastes, sports allegiances, recreational preferences) and the family and friends with whom we identify (Belk, 1987). Maffesoli (1996) has suggested that the contemporary equivalents of tribal affiliations are the "brand tribes" with which we identify (*see* BRAND COMMUNITY). Thus, we might identify ourselves as being a member of the "cult" of Macintosh, a "Red Devil" supporter of Manchester United Football Club, a Phish Phan, or a Harley Davidson motorcycle rider. We may also display these allegiances through

decals, signature clothing, and tattoos so that our affiliations are evident even in contexts where the focus of our loyalty is absent.

More subtly, we often express our moods, lifestyles, religions, personalities, political orientations, and much more by our choices of foods, clothing, homes, automobiles, and other visible consumption choices. There is a behavioral advantage in such attempts at nonverbal communication via consumption choices. Although it would be considered gauche, ostentatious, or crude to introduce yourself to other people by saying such things as "I am rich," "I am adventurous," "I've been to Bali," or "I am looking for love," such messages may be presented in more socially acceptable ways via our visible consumption choices. Nonverbal consumption choices are largely visual, but can also be olfactory, as with perfumes, incense, and cigarettes. Such material expressions of self are enacted largely through our possessions, but they can also involve such patronage choices in areas such as travel (e.g., first class, tour groups), clubs, restaurants, and retailers. And as we invest more of ourselves in digital representations and social networking sites (*see* SOCIAL NETWORKS), identity creation and expression can also include our photographs, web links, taste statements, virtual friendships, and other creative activities made visible to those who visit our Facebook page, website, or blog, as well as to those who receive our e-mails and text messages. With such means at our disposal, the lines of what is socially acceptable have been extended and more and more of our consumption and consumption-related aspirations have potentially become visible to others. Likewise, our network of other people who comprise a part of our extended self has also potentially become more readily apparent.

The Nature of the Self

One take on the contemporary self is that we are not only adapting a more malleable extended self that is increasingly external to our physical self and personal traits, but is also more postmodern, fragmented, decentered, and uncertain. Rather than a singular coherent core self that is consistent throughout our lives, in this view, we have multiple selves that we can put on and take off just as we change our wardrobe and clothing. Although the original conception of the extended self envisioned a central core self with various aggregate levels that are more or less central to our identity as well as various possessions that are also regarded as more proximate or distal to our sense of self, there are other possibilities.

One alternative conception of self that is consistent with the postmodern idea of fragmented selves is that we have a flexible cast of selves that we can call up depending on the situation and the others to whom we are presenting our selves. Thus we behave, dress, eat, and talk differently in a job interview, with our parents, with our children, with our significant other, and with our friends. In a study (Tian and Belk, 2005) of the office place possessions of employees at a high-tech firm in the United States, it was found that their offices, cubicles, and desks were a battleground where conflicting loyalties to the firm and their families were played out. Family members sought to keep themselves foremost in the employee's thoughts through photographs, mementos, children's artworks, telephone calls, e-mails, and other reminders and links. Meanwhile, the corporate culture of the firm made demands for offices uncluttered by such personal items and for some display of the corporate colors, especially when clients came to visit. In addition, the corporation's hold on employees was extended through expectations of employees putting in overtime hours as well as links to them outside of the office via telephone, e-mail, the Internet, pagers, and corporate events like picnics. Thus, rather than forming differing concentric levels of self, the familial and corporate selves engaged one another as adversaries competing for the employee/family member's time and devotion.

Another alternative to conceptualization of self that is more premodern than postmodern is when the nature of the self is more collective and inclusive rather than individual and exclusionary. For example, among the Iñupiaq of Alaska, it is believed that humans and whales share a kindred spirit (Bodenhorn, 2005). The whale gives itself to the people in order to sustain them, but its spirit does not die. Rather it returns again and again to feed the people when they are hungry. As Bodenhorn (2005) explains:

> What we are talking about here is neither the bounded individual nor performative fragmentary individuals who sound so postmodern. Instead we are talking about being part of a whole that is neither indistinguishable nor separate from it. The cosmology of the whale/human relationship is profoundly unifying (p. 91).

As Belk (1988) recognized, the concept of extended self is both Western and male. To the Iñupiaq sense, self is not an individualistic one radiating outward to incorporate more things and people in the environment. Rather, both the people and the environment are part of an integrated whole. The concept of individual ownership and possession is less relevant in this context and the alternative formulation of a giving holistic environment emphasizes the inseparability and interpenetration of people, animals, and the natural world.

The Nature of Possessions

The nature of possessions seems much more fixed than the existential question of who we are. Yet, notions of property, ownership, and the more general issue of our relationship with the world around us are all less stable and immutable than might initially appear to be the case. For example, just as the concept of self can be more shared and mutual rather than individual, so can the concept of ownership. Communal ownership is a case in point here. What European settlers in North and South America derisively called *Indian giving* was merely a case of Native Americans viewing the land and its resources as jointly owned, or more properly as unowned and unownable by anyone (e.g., Parry, 1986). Similar differences in notions of property and systems of property rights imposed by colonialists have eventually led to conflicts on every continent. Although planned communal societies based on common ownership have often failed and been dismissed as utopian, the vision of communities based on shared ownership has lasted thousands of years. On a smaller scale, the sharing that takes place within families has likely lasted hundreds of thousands of years (Belk, 2010). Infants depend upon parental sharing for their survival. Children typically do not need permission to eat from the family's food pantry, enter the family's home, or use most of the resources that the family owns. Nor, in using these pooled resources, do they incur a debt that must be specifically repaid. The family in its pooling and allocation of resources comes close to fulfilling Louis Blanc's prescription: from each according to his or her abilities and to each according to his or her needs.

This is not to say that the possessions of families are all held in common. What may have once been the family's joint participation in musical performance through singing, playing musical instruments, and dancing, has subsequently been commodified through subsequent iterations as written music, phonograph recordings, tape recordings, compact discs, music videos, and digital files. In the process, the former family possessions of radios, record players, VCR, and hi-fi systems have become privatized as his/her/their transistor radios, boom boxes, Walkman devices, iPods, and personal computer files. The same is true of many other former family-shared possessions and rituals including rooms, meals, televisions, and cars.

What is true in the family is true in society as well. On one hand, the rise of the Internet and the ability to give and receive digital files while losing nothing and expending little more than electricity has led to an economy in which the majority of music and films are now exchanged either freely or outside of authorized channels in the form of so-called pirated or counterfeit copies. This is truly an economy of sharing rather than proprietary ownership. The same is true of many other things that we regularly give away and consume thanks to the Internet: e-mails, instant messages, websites, bulletin boards, chat rooms, and blogs full of free advice, reviews, recipes, directions, maps, and so on. On the other hand, this free sharing of information has led to a rapid and expansive rise of intellectual property rights legislation designed to preserve the profits and proprietary ownership of music, films, genetic codes, software, and brand property, even (and especially) when it exists in the form of digital information (e.g., see Benkler, 2006). Mark Getty, chairman of Getty Images, has referred to intellectual property as the oil of the twenty-first century. No doubt there will be many more battles in the post-Napster post-Pirate Bay world of

file-sharing and fair-use skirmishes, but meanwhile entities like Google, Wikipedia, Linux, Flickr, Facebook, and others show how a nonproprietary sharing-based orientation toward information can still be an enormously successful business model. In an age of finite and dwindling resources, global warming, and vast inequalities between rich and poor, an economy of sharing looms as a highly enticing alternative to individualistic property and ownership.

Perhaps the world of academic journal publishing offers a model of how sharing, at least in the case of ideas and information, can be a win–win situation. The open model of science that has prevailed since the scientific revolution is based on the premise that science will advance much faster and all will benefit if we publish and share our ideas openly rather than seeking to profit from them. So we submit accounts of our research to journals for which we and others volunteer our reviewing expertise without compensation. We sign over the copyright of our work to the journals that agree to publish it. And we are happy when someone references our work and builds on our ideas. None of this needs be done out of a spirit of altruism and generosity, for when our work is published and cited, when we are listed as a member of an editorial review board or an editor, and when we send hard or digital copies of our work to anyone who seeks them, we are publicizing ourselves and building our academic reputations. But we have found a way to do this by giving away our ideas and services with the facilitating help of publishers, journals, and web sites who stand to earn a profit in the process. However, even here we can see the beginning of a more proprietary attitude toward the ownership of such ideas. Especially in the sciences such as genetic research, and in certain applied fields like electrical engineering, universities, laboratories, and corporations all seek to protect their inventions through patents, copyrights, and other forms of intellectual property. Scientists are reluctant to publish their work until they are certain they will obtain the patent rights to their discoveries. Thus, even within the hallowed halls of academia, the battle for property rights is increasingly in evidence.

Creating and Protecting an Intangible Possession: the Brand

Consumers are not the only ones concerned with creating self meanings through objects in the marketplace. Marketers are critically interested in investing meanings in their intangible brands. The two processes – marketers instilling meanings in their brands and consumers using these brands to create and maintain their self meanings – are interdependent. Moreover, since marketers can only do so much in their attempts to create brand meanings, the completion of this process ultimately depends to varying degrees on consumers to cocreate meanings for the brands they acquire in the marketplace. For example, the image of a model of Nike basketball shoe or Apple iPod depends partly on who the various people are who are first to adopt these objects, their characteristics such as the coolness that others ascribe to them, and what they have to say about their new acquisitions. Therefore, when a marketer attempts to measure the personality or image of a brand, they are partly measuring the image of the consumers who use the brand. And consumers for their part depend on the meanings that marketers have been able to impart to their brands via design, advertising, merchandizing, packaging, pricing, distribution, and other marketing activities. Celebrity endorsers often help in this process as well, as with the successes of Nike and Michael Jordan. It is easy to see this as a case of contagious magic in which the characteristics of the endorser transfer to the brand and from there hopefully to the purchaser of the goods being endorsed.

One category of goods where brand and user image is all-important is luxury goods. In fashions, accessories, handbags, luggage, fragrances, jewelry, and watches, luxury brands can sell for premium prices that can be as much as 10 times more expensive than average brands in these categories, or even more. Exclusivity, quality, and hip and elegant design are a part of the appeal of such goods, as is their exorbitant price and their sometimes limited availability. But their appeal also benefits from the image imparted through the use of celebrities and models promoting the brands, including Tiger

Woods, George Clooney, Chris Jorgensen, Uma Thurman, Kate Moss, Naomi Campbell, and Scarlett Johansson. Contagious magic is no less implicated in luxury goods than it is for more moderately priced offerings. In the case of Louis Vuitton, celebrities are also sought in the art world for design collaborations (e.g., Takeshi Murakami, Stephen Sprouse) and the company also sponsors art exhibitions featuring the work of these and other artists. Such cobranding seeks to draw on the high cultural capital of associating with the world of fine art; something that would be lost for example in sponsoring NASCAR race cars. Luxury goods companies like Tag Heuer that do seek sports celebrity endorsements confine themselves to higher status sports like golf and tennis rather than boxing, for example. Thus just as consumers can enhance their self image by surrounding themselves with popular goods and people, so can luxury brands enhance their images through associating themselves with high-status people, events, and settings.

Given the coproduction of brand meanings by consumers and marketers of luxury goods, one problematic occurrence is counterfeiting. In this case, the much lower prices of counterfeit goods make them available to more people and less-affluent people than comprise the target market of the authentic brands. This creates a dilemma. On one hand, such counterfeit brand proliferation threatens to dilute the brand and reduce its exclusivity. But, on the other hand, to have a luxury brand that is not the object of counterfeiting would suggest that the brand is not sufficiently desirable to warrant imitation. France, the home of many luxury brands including those of Louis Vuitton Moët Hennessy, has enacted legislation that makes it a crime to be in possession of a counterfeit good, with penalties of up to 3000 Euros per item, plus jail time.

On the other hand, brands like Nike, Stussy, Hilfiger, Polo, and DKNY all tolerated, if not encouraged, counterfeit copies of their brands among members of hip-hop culture in New York City. Their reason is simple. Rather than selling an exclusive luxury image, they are selling cool. By associating their brands with rap music and cool kids, they are trying to endow the goods with coolness. Product placements with popular cool kids in schools are another way that such brands sometimes try to achieve cool status. And another tactic is to use cool celebrities. For example, in a 2006 ad in *Transworld Skateboarding*, Vans promotes a skateboarding shoe called the *Hosoi SK8-HI*. The only copy reads "Very Limited Edition, Available April 1." The ad shows an action shot of skateboarding star Christian Hosoi, for whom the shoe is named. In the same way that Nike was able to exploit Michael Jordan's coolness with the Air Jordan, Vans hopes that Hosoi's cool image will transfer to the shoe. His cool image derives not only from his prominent role in early California skateboard culture, but also from his drug-dealing past. He was imprisoned for three years and was released shortly before the ad ran. He also appeared in the Vans-financed film, *Dogtown and Z-Boys*, about the origins of skateboarding in southern California. Vans has successfully used their early popularity in skateboarding subculture in order to borrow the cool renegade image of skateboarders. They have parlayed this image with products centered on music and cool sports including surfing, snowboarding, BMX wakeboarding, motocross, and supercross. Other shoes have tried to create a cool image by showing ads with cool cars and their owners (e.g., K-Swiss, 310), adult film stars (e.g., Pony), or simply cool-looking people with tattoos and a look of cool indifference (e.g., GBX).

So it can be seen in these examples that the implications of possessions for self image can be commercial as well as personal. By focusing on extended self, celebrity endorsement, cool people, cool settings, and contagious magic, it is sometimes possible to build a brand into an aspirational prop for self image. If William Shakespeare and Erving Goffman are correct, all the world is a stage, and we increasingly rely on ensembles of branded consumer goods in order to try on the various roles that we imagine ourselves playing.

Conclusion

It is easy to condemn our reliance on possessions to enlarge our sense of self as being a shallow substitute for the alternative character-based sense of self that presumably hides behind our

façade of consumer goods. Materialism, waste, excessive consumption, and the vicious cycle of striving to have the latest gadgets ultimately prove unfulfilling. And there is much to be said for striving for a simpler material lifestyle and reducing our ecological footprint, especially among those of us who have much more than we need in order to live comfortably. But despite these easy criticisms, possessions also play some very positive roles in our lives. By accumulating certain possessions, we are accumulating visual mnemonic cues to our pasts and potential cues to the meaning of our lives. These material cues often take the form of photographs, gifts, souvenirs, awards, and heirlooms, but they can as easily be a fragment of a song, the whiff or a perfume, or the taste of a particular food like Proust's lime-blossom tea and petite madeleines. At more aggregate levels of self, we also have national monuments, national buildings, flags, and other icons to remind us of our country's past.

Besides looking backward and reminding ourselves of who we are, possessions that we aspire to own can also provide goals that drive us forward in education, career, personal relationships, travel, and other spheres of life. Aesthetic objects can be a source of contemplation, relaxation, and reflection about who we are and what is important in life. We have come to rely on material goods in our rituals of birth, death, marriage, childbirth, and other rites of passage that help us get through life and feel that we are honoring life's sacred moments. Even in our so-called disposable age of ephemeral possessions, it is likely that some of our individual and collective possessions will outlive us. They are sometimes our link to immortality and parts of the trace that we leave behind that will help others remember us when we are gone. For academics, artists, writers, playwrights, poets, filmmakers, craftspeople, builders, cabinetmakers, and others who create, it is their creations that they hope will fulfill this function. For others, it may be any of the many goods that have become embedded in their lives and identities. But regardless of how humble or magnificent the object, the power of possessions lies not only in their role in our present lives, but potentially in the past and future as well.

Bibliography

Belk, R. (1985) Materialism: trait aspects of living in the material world. *Journal of Consumer Research*, **12** (3), 265–280.

Belk, R.W. (1987) Identity and the relevance of market, personal, and community objects, in *Marketing and Semiotics: New Directions in the Study of Signs for Sale* (ed J. Umiker-Sebeok), Mouton de Gruyter, Berlin, pp. 151–164.

Belk, R.W. (1988) Possessions and the extended self. *Journal of Consumer Research*, **15** (2), 139–168.

Belk, R.W. (1992) Attachment to possessions, in *Human Behavior and Environment: Advances in Theory and Research, Vol. 12, Place Attachment* (eds I. Altman and S., Low), Plenum Press, New York, pp. 37–62.

Belk, R.W. (2010) Sharing. *Journal of Consumer Research.* **34** (1), 1–20.

Benkler, Y. (2006) *The Wealth of Networks: How Social Production Transforms Markets and Freedom*, Yale University Press, New Haven, CT.

Bodenhorn, B. (2005) Sharing costs: an exploration of personal and individual property, equalities and differentiation, in *Property and Equality, Vol. 1: Ritualisation, Sharing, Egalitarianism* (eds T. Widlock and W.G. Tadesse.), Berghahn Books, New York, pp. 77–104.

Csikszentmihalyi, M. and Rochberg-Halton, E. (1981) *The Meaning of Things: Domestic Symbols and the Self*, University of Chicago Press, Chicago.

James, W. (1890) *The Principles of Psychology, Vol. 1*, Henry Holt, New York.

Kleine, S.S. and Baker, S.M. (2004) An integrative review of material possession attachment. *Academy of Marketing Science Review*, **1**, 1–39.

Maffesoli, M. (1996) *The Time of the Tribes: The Decline of Individualism in Mass Society*, D. Smith, trans., Sage, Thousand Oaks, CA.

Parry, J. (1986) The Gift, the indian gift and the 'Indian Gift'. *Man*, **21**, 466–471.

Price, L.L., Arnould, E.J., and Curasi, C.F. (2000) Older consumers' disposition of special possessions. *Journal of Consumer Research*, **27**, 179–201.

Tian, K. and Belk, R.W. (2005) Extended self and possessions in the workplace. *Journal of Consumer Research*, **32**, 297–310.

Wicklund, R. and Gollwitzer, P. (1992) *Symbolic Self-Completion*, Lawrence Erlbaum, Hillsdale, NJ.

attitudes

Paul M. Herr

Any comprehensive treatment of the attitude construct in general and its applications in marketing in particular could easily fill this volume, let alone a single article. Hence, the focus here is to simply identify the construct, describe some of the major stages in its development, and briefly trace its recent history in consumer behavior. Consequently, this article will be far more selective than comprehensive in nature.

Early Development

Throughout this discussion, one is struck by the singular contributions of attitude theorists in social psychology to applications in consumer research. Attitude has long been viewed as a central construct in social psychology, and its role in consumption has been viewed as no less important. Clearly, the ability to elicit from consumers a simple, summary evaluation of goods and services with predictive utility has had strong allure. This appeal, however, is tempered by a history rich with disappointment and debate. Even something as (deceptively) simple as defining the attitude construct (which could also easily fill this volume), not to mention settling debate over the construct's very existence, has filled academic journal pages for many decades. Thurstone's (1928) triumphant claim that attitudes can be measured presupposed both their stable existence and their directive influence on an individual's behavior. Allport's (1935) early definition of attitude continues to guide current thinking about attitude's functional role, as well as spur debate. According to Allport, "An attitude is a mental and neural state of readiness, organized through experience, exerting a directive or dynamic influence upon the individual's response to all objects and situations with which it is related" (1935, p. 810). Central to this definition is attitude's enduring nature, that it is learned through experience, and its directive influence on behavior.

Attitudes' functional role was considered by Katz (1960). His analysis provides compelling reasons why consumers may hold attitudes as well. Katz suggested that individuals might be well-served by holding stable attitudes about persons and objects in their environment, as useful functions might be served by attitudes. Specifically, Katz identified four functions that attitudes might provide. It is a simple extension to note what each provides the consumers. First, consumer attitudes may serve a *utilitarian* function. That is order and structure are provided by knowing which brands fill needs well and which do not. Likewise, a *value expressive* function is served to the extent a consumer is able to identify which brands and products allow them to best express their central values and who they believe themselves to be. Attitudes of both valences about luxury goods may serve this function. An *ego-defense* function is served by attitudes that protect against damage to the self-concept. Finally, a *knowledge* function allows the holder of attitudes to make sense of an otherwise chaotic and disorganized world; knowing which brands are loved and which are hated.

The Theory of Reasoned Action

The existence of functionally sound reasons to hold attitudes notwithstanding, prominent failures of attitudes to in fact predict behaviors led Wicker (1969) to suggest that attitudes are epiphenomenal; after-the-fact explanations of behavior, rather than guiding influences, and as such should be abandoned as a scientific construct. Two responses to this attack on attitudes are particularly noteworthy, with respect to both their impact within traditional attitude theory and their influence on marketing. First, with their Theory of Reasoned Action Fishbein and Ajzen (1975) attempted to salvage the attitude construct by (i) redefining attitude as an expected value of the sum of an individual's beliefs about an attitude-object, weighted by individual's evaluations of each of those beliefs, (ii) considering the impact of subjective norms for engaging in attitude-relevant behavior, and (iii) weighting the attitude and subjective norms in predicting the individual's intention to engage in a particular behavior. This development's impact on marketing has been profound. The bulk of academic marketing work, as well as marketing practitioners' work

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

in attitudes, employs this framework or variants (and more recently, the Theory of Planned Behavior, Ajzen, 1985).

Academic research explicitly studying consumers generally has taken one of two paths: either directly applying the model to consumption settings (and more recently, applying the models to social marketing settings) or engaging in theoretical discussions of the model itself. Representative of the former research is work by Shimp and Kavas (1984) investigating consumer coupon usage and Hansen, Jensen, and Solgaard (2004) examining online grocery purchase intentions.

A great deal of work has investigated modified versions of the Theory of Reasoned Action, as well as questioned the assumptions of and made extensions to the original model. Representative work includes Bagozzi (1981, 1982), Sheppard, Hartwick, and Warshaw (1988), Miniard and Cohen (1983), and Mitchell and Olsen (1981). Central to this work are concerns for the ability of attitudes (affective judgments) to be fully mediated by expectancy-value judgments, the process by which expectancy-value judgments are expected to influence behavior, and how the entire Theory of Reasoned Action model's components is expected to influence behavior.

The Process Model of Attitude-Behavior Consistency

The second significant response to Wicker's (1969) assault on attitudes was a focus on *how* attitudes may influence behavior (Fazio and Zanna, 1981; Fazio *et al.*, 1986). This latter development shifted focus away from "Do attitudes predict behavior?" to a process question "How do attitudes guide behavior?" Moreover, the definition of attitude proposed in this work is conceptually closer to Allport's definition of four decades earlier, but without the necessity that attitudes be linked to behavior. Rather, that link's existence was left as an empirical question. An attitude is considered to be a learned association between a concept and an evaluation (e.g., an affective categorization of a person or an object). Structurally, this may be represented in memory as a link between a node representing the object or person, and its evaluation (e.g., good, bad, and likable.). Operationally, attitudes may vary in strength, reflecting the strength of the associative link between the object and its evaluation. Relatively weak links reflect relatively weak (or nonexistent) attitudes. Encountering the attitude-object (or its representation) is insufficient to activate the associated evaluation from memory, so the attitude fails to influence behavior directed toward the person or object. Relatively strong links, however, reflect strong attitudes. Simply encountering the attitude-object (or its representation) may automatically activate the individual's evaluation of the person or object from memory. Once activated, the attitude serves as a filter, through which the object and situation are perceived, filtering out attitude-inconsistent information and coloring ambiguous information attitudinally consistent. The individual's behavior is determined by their immediate perceptions of the situation (biased by the attitudinal filter). Attitude-behavior consistency occurs not because of a striving for consistency, or as an end result of a conscious, rational process but rather as a natural end state of the attitude's activation (of which the attitude holder may be wholly unaware). Moreover, for strong attitudes, the activation itself may be automatic and wholly outside the control of the individual (Fazio *et al.*, 1986).

Applications of Fazio's model in marketing have been relatively few, but its applicability is well demonstrated by Fazio, Powell, and Williams (1989). The authors demonstrated that, even after controlling for attitude extremity, product choice was influenced by attitude accessibility. Branded products for which participants held relatively accessible attitudes were more likely to be chosen than products for whom participants held less accessible attitudes. Moreover, prominently positioned products (those placed in the front row of a display) were more likely to be chosen by participants holding less accessible attitudes. That is, their choice was determined by a situational factor (salience of position) rather than an attitudinal factor (strength).

To a considerable extent, attitude's definition has determined the types of models employed in predicting attitude's influence on behavior. While debate over attitude's definition appeared

to be subsiding in the last decade of the twentieth century, the distinction between implicit and explicit attitudes has revived interest in definitional matters, and the debate has been rejoined. This distinction will be visited below.

Attitudes and Persuasion

A central task facing consumer researchers is to understand the conditions under which product and service communications will produce a sufficiently strong attitude to result in consumer choice favoring a given product or service. During the 1980s considerable research addressed exactly this issue, growing out of Chaiken's (1980) Heuristic-Systematic Persuasion model and Petty and Cacioppo's (1986) Elaboration Likelihood models. The models make largely similar predictions. The ELM (Elaboration Likelihood Model) identifies two "routes" to persuasion; central and peripheral. The central route involves cognitively effortful consideration of the arguments (content) of the communication message. Central processing occurs only when the individual is both motivated and able to think about the message and product. If either is lacking (e.g., the individual does not care about the message or product due to irrelevance, or is distracted, or otherwise lacks the cognitive resources to process the message) the central processing route will be aborted in favor of the peripheral route.

Message-consistent attitudes growing out of the peripheral route to persuasion are based not on the strength of the message arguments, but rather on peripheral cues to message strength, that may be wholly invalid with respect to indicating actual strength of the message. For instance, cues such as number of arguments, length of the argument, communicator attractiveness and confidence all signal that the message may be credible. Without analyzing the actual content of the message, though, a judgment of message credibility is premature. Followers of the peripheral route are insufficiently motivated, or lacking in ability (or both) to effortfully process the content of the message and instead simply rely on the peripheral cues when forming their communication-based attitude.

Central to the ELM is the notion that attitudes changed or formed via the central route will have different effects than will attitudes changed through the peripheral route. Specifically, central route-based attitudes show greater temporal persistence, greater resistance to change, and are better predictors of behavior than attitudes that result mostly from peripheral cues.

The ELM and heuristic-systematic model (HSM) have garnered considerable empirical support, within both the field of consumer behavior and a far-ranging variety of other areas. In a consumer setting, Petty, Cacioppo, and Schumann (1983) exposed participants to print advertisements for a new razor blade. Whether the central or peripheral route would be followed was manipulated by the relevance of the ad to readers. High issue relevance (involvement) was induced by indicating that the razor would be available for purchase in their local market area. It was expected (and confirmed) that this would lead to greater reliance on central processing. Low issue relevance was induced by indicating to participants that the advertised razor would not be available in their market area. Under low relevance conditions, individual's attitudes were most affected by the celebrity status of the product endorsers. In contrast, the quality of product arguments was a more important determinant of the attitudes of individuals exposed to the advertisements under high relevance conditions.

While the ELM and HSM deal primarily with persuasion (attitude change) Fazio's (1990) Mode model addresses how attitudes influence behavior, and reconciles the Theory of Reasoned Action with his own process model. A key distinction between the two models involves the amount of cognitive processing involved in deciding how one will behave in the presence of the attitude object. The Theory of Reasoned Action presents behavior as a reasoned, cognitively cumbersome, and highly deliberative process on the part of the individual. Individuals are said to make systematic use of the information at hand. Salient beliefs about the attitude-object are weighted by their evaluations, summed, and compared with norms to form a behavioral intention. The individual is expected to behave consistently with the intention. Fazio's process

model presents an alternative in which activation of an attitude is relatively effortless and automatic. The attitude colors immediate perceptions of the object in the current situation, which in turn may prompt attitudinally congruent behavior. The MODE model suggests that both attitude-behavior processes may occur, but the antecedents of each differ. Specifically, the more cognitively effortful process occurs only when the individual is motivated and has the opportunity to engage in the more taxing process. If either antecedent is absent, the default value is the relatively effortless process of the Fazio model. In the realm of consumer behavior, the expected differences in attitude-to-behavior processes are relatively straightforward. When consumers are highly involved and have the ability to process, they will likely engage in reasoned action. Such cases include big-ticket items, or purchases of considerable significance in some other regard. Frequent purchases or purchases with small downside potential seem more likely to be driven by the relatively effortless, top-down guidance provided by a highly accessible attitude.

Implicit Attitudes

A relatively recent development of some significance is the notion of implicit attitudes and the development by Greenwald and his colleagues of the Implicit Association Test (IAT). Greenwald and Banaji (1995) define *implicit attitude* as "... introspectively unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feeling, thought, or action toward social objects." (p. 8). This definition grew out of the work in social psychology examining "automatic" (as used by Bargh and his colleagues, e.g. Bargh, 1994) behavior, as well as Fazio's definition of attitude as an association (of varying strength) between an object and its evaluation. "Automatic" behavior is said to proceed in an implicit or unconscious fashion. It is characterized by its spontaneous, unplanned nature, over which the actor has neither awareness nor control, nor, most importantly, awareness of (or conscious access to) cognitions (including attitudes) causally precipitating the behavior. Behavior of this type is at an extreme end of the planned-spontaneous behavior continuum.

The implicit nature of attitudes (and their measurement) is interesting and important, according to Nosek, Greenwald, and Banaji (2007) because, "... implicit cognition could reveal traces of past experience that people might explicitly reject because it conflicts with values or beliefs, or might avoid revealing because the expression could have negative social consequences. Even more likely, implicit cognition can reveal information that is not available to introspective access even if people were motivated to retrieve and express it (see Wilson, Lindsey, and Schooler, 2000, for a similar theoretical distinction for the attitude construct specifically) .... The term *implicit* has come to be applied to measurement methods that avoid requiring introspective access, decrease the mental control available to produce the response, reduce the role of conscious intention, and reduce the role of self-reflective, deliberative processes" (p. 266).

Greenwald and his colleagues have developed a measure of implicit attitudes (the IAT) that is both widely available and easily used. A description of the test from his website follows:

"In the IAT a subject responds to a series of items that are to be classified into four categories – typically, two representing a concept discrimination such as *flowers* versus *insects* and two representing an attribute discrimination such as *pleasant* versus *unpleasant* valence. Subjects are asked to respond rapidly with a right-hand key press to items representing one concept and one attribute (e.g., *insect*s and *pleasant*), and with a left-hand key press to items from the remaining two categories (e.g., *flowers* and *unpleasant*). Subjects then perform a second task in which the key assignments for one of the pairs is switched (such that *flowers* and *pleasant* share a response, likewise *insects* and *unpleasant*). The IAT produces measures derived from latencies of responses to these two tasks. These measures are interpreted in terms of association strengths by assuming that subjects respond more rapidly when the concept and attribute mapped onto the same response are strongly associated (e.g., *flowers* and *pleasant*) than when they are weakly associated (e.g., *insects* and *pleasant*)." (http://faculty.washington.edu/agg/iat_materials.htm)

The IAT measure is not without its critics. Blanton and Jaccard (2008) note that many variables are confounded with associative strength, and influence the IAT, including the similarity between two objects (De Houwer, Geldof, and De Bruycker, 2005), different degrees of familiarity with stimulus items on the IAT (Brendl, Markman, and Messner, 2001), and "salience asymmetry," of items on the IAT (Rothermund and Wentura, 2004). Reliability concerns with the IAT have also been raised. Steffens and Buchner (2003) found test-retest correlations ranging from 0.50 to 0.62. Greenwald, Nosek, and Sriram (2006) reported average test-retest scores of the IAT across a range of studies of 0.56. Cunningham, Preacher, and Banaji (2001) reported average test-retest correlations over a two week measurement period of 0.27.

Moreover, Han *et al.* (2010) found that the IAT's susceptibility to extrapersonal (essentially nonattitudinal) associations increases its sensitivity to transient contextual influences (that render both reliability and validity suspect). Specifically, the authors compared the IAT with Olson and Fazio's Personalized IAT (in which the evaluative labels include "I" (e.g. "I like" or "I dislike") rather than a general "Like" or "Dislike." This forces respondents to interpret the stimuli in terms of their own experience, rather than the ambiguous interpretation (either personal or normative) that is possible when a nonspecific label (such as "good") is used. In the first experiment, the traditional IAT was found susceptible to mindset priming effects. An earlier experience in an unrelated task influenced how respondents disambiguated the IAT labels, and whether they adopted a normative or a personal focus while completing the IAT. Subsequent experiments revealed that scores on the IAT indicated attitude change when it was unlikely to have occurred, and failed to detect attitude change in a situation when all other indications were that attitude change had in fact occurred. In each case, by personalizing the IAT the influence of contextual factors was eliminated.

In spite of predating the formal development of the implicit attitude construct, the MODE model anticipates its relevance and accommodates its existence in consideration of attitude function surprisingly well. This is partly due to the definition of attitude in both cases as dependent on associative strength, but also due to the MODE model's explicit accounting for how attitudes are expected to influence behavior. That is, a person's awareness of the attitude (or its basis) may have little to do with whether the attitude can influence behavior. As discussed previously, of greater import is the strength of the attitude and the motivation and opportunity of the individual to process. When the association between the object and its evaluation is sufficiently strong, evaluations may be activated automatically upon the encounter of an object-relevant stimulus.

The activated attitude's influence on further evaluative judgments of the attitude-object depends on the individual's motivation and opportunity to engage in elaborate processing of evaluative characteristics of the attitude-object.

Fazio (2007) notes that the correspondence between implicit and explicit measures of attitudes also depends on the motivation and opportunity of the individual. Specifically, he suggests that since the verbal expression of any object-judgment occurs following the automatic activation of any relevant attitude, such attitude-relevant verbal responses may be influenced by motivational factors as well. Hence, greater correspondence between implicit and explicit measures is expected when motivation to deliberate is low and (or) the opportunity to do so is low. When both are high, additional information is likely to be at least to some extent inconsistent with the attitude indicated by the implicit measure, rendering correlation low.

Conclusion

Albeit well-intentioned and within constrained limits successful, the quest for a simple measure of attitudes' influence on behavior appears all but finished, and for those interested in simple answers, the news is not good. As noted, though, the journey has provided insight into a far more interesting set of questions, and uncovered relations between attitudinal processing and behavior that are both more interesting and important for a far broader range of attitudinal phenomena than the answer to the simpler "Does attitude predict behavior?" question. Specifically, attitude existence, verbal expressions of, formation, strength, change,

resistance to attack, persuasion, conditions under which some kinds of attitudes predict some kinds of behavior for some kinds of people, and attitudes' relations to other constructs have all become substantial areas of investigation in their own right. Each also has much to inform the student of consumer behavior.

See also *attitude–behavior consistency; consumer categorization; consumer decision making; consumer intentions; consumer involvement; consumer memory processes; implicit consumer cognition; knowledge accessibility; persuasion*

Bibliography

Ajzen, I. (1985) From intentions to actions: a theory of planned behavior, in *Action-control: From Cognition to Behavior* (eds J. Kuhl and J. Beckman), Springer, Heidelberg, pp. 11–39.

Allport, G. (1935) Attitudes, in *A Handbook of Social Psychology* (ed. C.M. Murchison), Clark University Press, Worchester, MA, pp. 798–844.

Bagozzi, R.P. (1981) Attitudes, intentions, and behavior: a test of some key hypotheses. *Journal of Personality and Social Psychology*, **41**, 607–627.

Bagozzi, R.P. (1982) A field investigation of causal relations among cognitions, affect, intentions, and behavior. *Journal of Marketing Research*, **19** (4), 562–583.

Bargh, J.A. (1994) The four horsemen of automaticity: awareness, intention, efficiency, and control in social cognition, in *Handbook of Social Cognition*, Vol. 1, 2nd edn (eds R.S. Wyer Jr., and T.K. Srull), Erlbaum, Hillsdale, NJ, pp. 1–40.

Blanton, H. and Jaccard, J. (2008) Unconscious racism: a concept in pursuit of a measure. *Annual Review of Sociology*, **34**, 277–2097.

Brendl, C.M., Markman, A.B. and Messner, C. (2001) How do indirect measures of evaluation work? Evaluating the inference of prejudice in the implicit association test. *Journal of Personality and Social Psychology*, **81**, 760–773.

Chaiken, S. (1980) Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, **39**, 752–756.

Chaiken, S. (1987) The heuristic model of persuasion, in *Social Influence: The Ontario Symposium*, Vol. 5 (eds M.P. Zanna, J.M. Olson and C.P. Herman), Erlbaum, Hillsdale, NJ, pp. 3–39.

Cunningham, W.A., Preacher, K.J. and Banaji, M.R. (2001) Implicit attitude measures: consistency, stability, and convergent validity. *Psychological Science*, **12**, 163–170.

De Houwer, J., Geldof, T. and De Bruycker, E. (2005) The implicit association test as a general measure of similarity. *Canadian Journal of Experimental Psychology*, **59**, 228–239.

Fazio, R.H. (1990) Multiple processes by which attitudes guide behavior: the MODE model as an integrative framework, in *Advances in Experimental Social Psychology*, Vol. 23 (ed. M.P. Zanna), Academic Press, San Diego, pp. 75–109.

Fazio, R.H. (2007) Attitudes as object-evaluation associations of varying strength. *Social Cognition*, **25**, 603–637.

Fazio, R.H., Powell, M.C. and Williams, C.J. (1989) The role of attitude accessibility in the attitude-to behavior process. *Journal of Consumer Research*, **16** (3), 280–288.

Fazio, R.H., Sanbonmatsu, D.M., Powell, M.C. and Kardes, F.R. (1986) On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, **50**, 229–238.

Fazio, R.H. and Zanna, M.P. (1981) Direct experience and attitude-behavior consistency, in *Advances in Experimental Social Psychology*, Vol. 14 (ed. L. Berkowitz), Academic Press, New York, pp. 161–202.

Fishbein, M. and Ajzen, I. (1975) *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*, Addison-Wesley, Reading, MA.

Greenwald, A.G. and Banaji, M.R. (1995) Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review*, **102**, 4–27.

Greenwald, A.G., Nosek, B.A. and Sriram, N. (2006) Consequential validity of the implicit association test: comment on Blanton and Jaccard. *American Psychologist*, **61**, 56–61.

Han, H.A., Czellar, S., Olson, M.A. and Fazio, R.H. (2010) Malleability of attitudes or malleability of the IAT? *Journal of Experimental Social Psychology*, **46** (2), 286–298.

Hansen, T., Jensen, J.M., and Solgaard, H.S. (2004) Predicting online grocery buying intention: a comparison of the theory of reasoned action and the theory of planned behavior. *International Journal of Information Management*, **24**, 539–550.

Katz, D. (1960) The functional approach to the study of attitudes. *Public Opinion Quarterly*, **24**, 163–204.

Miniard, P.W. and Cohen, J.B. (1983) Modeling personal and normative influences on behavior, *The Journal of Consumer Research*, **10**, 169–180.

Mitchell, A.A. and Olson, J.C. (1981) Are product attribute beliefs the only mediator of advertising effects on brand attitudes? *Journal of Marketing Research*, **18** (3), 318–332.

Nosek, B.A., Greenwald, A.G. and Banaji, M.R. (2007) The implicit association test at age 7: a methodological and conceptual review, in *Automatic Processes in Social Thinking and Behavior* (ed. J.A. Bargh), Psychology Press, pp. 265–292.

Olson, M.A. and Fazio, R.H. (2004) Reducing the influence of extrapersonal associations on the implicit association test: personalizing the IAT. *Journal of Personality and Social Psychology*, **86**, 653–667.

Petty, R.E. (2001) Attitude change: psychological, in *International Encyclopedia of the Social & Behavioral Sciences* (eds J.S. Neil and P.B. Baltes), Pergamon, Oxford.

Petty, R. and Cacioppo, J.T. (1986) The elaboration likelihood model of persuasion, in *Advances in Experimental Social Psychology* Vol. 19 (ed. L. Berkowitz), Academic Press, New York, pp. 123–205.

Petty, R.E., Cacioppo, J.T. and Schumann, D. (1983) Central and peripheral routes to advertising effectiveness: the moderating role of involvement. *Journal of Consumer Research*, **10**, 135–146.

Rothermund, K. and Wentura, D. (2004) Underlying processes in the implicit association test: dissociating salience from associations. *Journal of Experimental Psychology*, **133**, 139–165.

Sheppard, B.H., Hartwick, J. and Warshaw, P.R. (1988) The theory of reasoned action: a meta-analysis of past research with recommendations for modifications and future research. *Journal of Consumer Research*, **15**, pp. 325–343.

Shimp, T.A. and Kavas, A. (1984) The theory of reasoned action applied to coupon usage. *Journal of Consumer Research*, **11**, 795–809.

Steffens, M.C. and Buchner, A. (2003) Implicit association test: separating transsituationally stable and variable components of attitudes toward gay men. *Experimental Psychology*, **50**, 33–48.

Thurstone, L.L. (1928) Attitudes can be measured. *American Journal of Sociology*, **33**, 529–554.

Wicker, A.W. (1969) Attitudes versus actions: the relationships of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, **25**, 41–78.

Wilkie, W.L. and Pessemier, E.A. (1973) Issues in marketing's use of multi-attribute attitude models. *Journal of Marketing Research*, **10**, 428–441.

Wilson, T.D., Lindsey, S. and Schooler, T. (2000) A model of dual attitudes. *Psychological Review*, **107**, 101–126.

implicit consumer cognition

Patrick T. Vargas and Sangdo Oh

Advertisers worry about clutter; they worry about getting their message noticed amid all the other messages aimed at consumers. Some estimates suggest that the average American is exposed to over 3000 advertisements each day. It would be impossible for even the most motivated and patient consumer to wade through all of that information attentively. Fortunately for advertisers, focused attention may not be absolutely necessary. Consumers may be affected by information outside of their conscious awareness. There is a large and growing body of psychological and consumer research demonstrating that people need not intentionally process information to be affected by it. In this article, we offer a brief review of research on implicit consumer cognition.

Implicit cognition (*see* KNOWLEDGE ACCESSIBILITY), more generally, refers to the influence of past experience without conscious recollection or awareness of the influencing experience. For example, consider a classic study of memory (Warrington and Weiskrantz, 1970). Amnesiacs have difficulty with short-term memory. When shown a list of words for study and later asked to recall words from the list, amnesiacs perform substantially worse than non-amnesiacs. However, on certain types of tests, amnesiacs show roughly the same evidence of prior exposure to the studied words as non-amnesiacs. These are implicit tests, in which the memory "test" is presented as a simple task to be completed rather than an explicit demand for conscious recollection, for example, word stem completion: deter_ _ _ _ could be either detergent or deterrent. Amnesiacs and non-amnesiacs that had recently studied a list of words including the word detergent are more likely to use detergent to complete the stem. Thus, amnesiacs show evidence of prior exposure to the words, at almost exactly the same rate as non-amnesiacs, but only on implicit measures of memory.

Implicit consumer cognition may be defined as the unacknowledged or misidentified influence of prior experience on consumer-related judgment and behavior. In this article, we review research on implicit consumer cognition. We begin with an overview of implicit measures, with a focus on how they may be applied to marketing and consumer behavior; and then we address implicit processes and effects, including topics such as subliminal persuasion and priming.

Implicit Measures of Attitudes

ATTITUDES are central to the study of marketing and consumer behavior. Consumers rarely choose products or patronize services that they dislike, so it is important for marketers to engender positive attitudes. Advertising agency Saatchi & Saatchi conceptualizes brand strength in terms of "lovemarks," employing ideas like love, respect, inspiration, sensuality, and intimacy. Very often consumers report definitive attitudes toward brands, sometimes positive and sometimes negative. Traditionally, attitudes are measured using explicit, self-report measures, in which respondents indicate on, say 10-point scales, how much they like or dislike an object (semantic differential scales). Other explicit attitude measures present respondents with a series of valenced statements about an object, and respondents are asked to indicate their level of agreement with the statements (Likert-type scales), or select a few statements that most closely match their own feelings (Thurstone scales). However, sometimes research participants may be unwilling to express their true feelings, or, in some cases, they may not know how they really feel about some object. In these cases, researchers have turned to implicit measures of attitudes.

Like implicit measures of memory (described briefly, above), implicit measures of attitudes are presented to research respondents simply as tasks to be completed. Respondents' performance on the task can then be used to infer their attitudes. In one early consumer study involving an implicit measure, respondents were simply given a seven-item grocery shopping list and asked to make some judgments about the woman who made the list (Haire, 1950). There were actually two versions of the list: one version contained Maxwell House drip coffee, and the other version contained Nescafe instant coffee. The lists were identical in all other respects. Half of the research participants received the Maxwell House version

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

of the list and the other half received the Nescafe version of the list. Those who received the Nescafe version were more likely to describe the woman who wrote the list as "lazy," "a poor planner," and a "bad wife." These characteristics strongly suggest that consumers had negative attitudes toward Nescafe instant coffee. Implicit measures of attitudes have come a long way since 1950.

Two of the most popular contemporary measures are computer-based and rely on millisecond differences in response times. The Implicit Association Test (IAT, *see* Greenwald *et al.*, 2009) is presented to respondents as a series of classification tasks. Four categories of stimuli are presented sequentially on a computer screen, and respondents must categorize the stimuli as quickly as possible by pressing one of two keys associated with the different categories. Initial trials of a typical IAT involve two categories, for example, good and bad, or fruit and cigarettes. Respondents must press one key if the stimulus is a good word or picture, and another key if the stimulus is a bad word or picture. Some subsequent trials involve all four categories, forcing respondents to pair categories on the keys. For example, one block of trials may require respondents to press one key if the stimulus is either fruit-related or good, and another key if the stimulus is either cigarette-related or bad. A later block of trials will reverse the pairings, requiring respondents to press one key if the stimulus is either cigarette-related or good, and another key if the stimulus is either fruit-related or bad. In this example, we refer to these blocks of trials as consistent (fruit/good and cigarettes/bad) and inconsistent (cigarettes/good and fruit/bad). For most Americans, categorizing the consistent trials would be easier, and faster, than categorizing the inconsistent trials. Average response times for each block of trials can be calculated, and the difference between these times is known as the *IAT effect*. It is really a measure of the strength of an individual's associations between the concepts used in the task. If someone finds it easier to associate fruit/good and cigarettes/bad than cigarettes/good and fruit/bad, we can infer that they have relatively favorable attitudes toward fruit, and relatively negative attitudes toward cigarettes.

The IAT has generated an enormous amount of research. A recent review of the research considered 184 distinct studies involving a total of 14,900 respondents (Greenwald *et al.*, 2009). A meta-analysis of these studies revealed that the IAT was moderately correlated with a wide variety of different behaviors, judgments, and physiological measures ($r = 0.274$). In the domain of consumer preferences (40 studies involving 3257 respondents), the IAT showed a slightly higher correlation with relevant outcomes ($r = 0.323$). Studies of political preference (11 studies, 2903 respondents) obtained average IAT-criterion correlations of 0.483; studies of alcohol and drug use (16 studies, 1718 respondents) obtained average IAT-criterion correlations of 0.221.

A second popular, computer-mediated implicit measure of attitudes relies on priming (*see* Wittenbrink, 2007). Measuring attitudes with priming also involves the sequential presentation of stimuli on the screen, and requires respondents to categorize the stimuli as quickly and accurately as possible. Again, respondents categorize the stimuli using two keys, each assigned to one of the two categories – usually one key for "good" and one key for "bad." The judgment task is influenced by the brief presentation of an attitude object immediately prior to the task, a prime. Reusing the example from above, a prime (picture of a cigarette or picture of a piece of fruit) would appear very briefly on the screen, and disappear, replaced with a good or bad word. The respondent's task is to categorize that target word. In general, people are faster to categorize the target words when they have been preceded by an evaluatively consistent prime. So, for example, people would be faster to categorize a good word when it is preceded by an image of grapes, than to categorize a bad word when preceded by an image of grapes. Likewise, they would be faster to categorize a bad word when a picture of a cigarette precedes it, than to categorize a good word when a picture of a cigarette precedes it. However, someone who loves cigarettes and hates fruit would likely be faster to categorize cigarette → good and grapes → bad, compared to grapes → good and cigarette → bad. This priming measure has also been used in a wide variety of studies, and

reliably predicted relevant behaviors and judgments. This measure has been shown to work when the primes are presented subliminally – so the procedure can be administered to respondents who are completely unaware of the attitude objects under consideration.

There are dozens of implicit measures in use now, each with different characteristics. Many of these implicit measures do not require millisecond-resolution for response times, and may therefore be easier to use in the absence of dedicated computer software. One such measure is the Affect Misattribution Procedure (AMP; Payne *et al.*, 2005). This measure is procedurally similar to the priming measure, but does not require respondents to categorize target stimuli. In the AMP, respondents are presented with a prime, that is, some attitude object (e.g., grapes or cigarettes), which then disappears from the screen and is replaced by some neutral, abstract stimulus (e.g., a Chinese ideograph). Respondents then have to indicate how much they like the ideograph. Their responses are influenced by the primes. Ideographs that are preceded by pleasant primes are rated as more favorable than those that are preceded by unpleasant primes. This effect occurs even when respondents are explicitly warned that the primes could influence their judgments, and instructed to try to prevent that from happening.

Other implicit measures are even easier to implement, requiring only willing respondents, a pencil, and some paper (*see* Vargas, Sekaquaptewa, and von Hippel, 2007). These measures come in many different forms, and have been used to predict judgments and behaviors, as well. One such measure involves presenting respondents with sentence beginnings, and simply asking respondents to complete the sentence in any way that is grammatically correct. Some of the sentence beginnings are constructed so that they violate or confirm commonly held beliefs, like the belief that Toyota automobiles are very reliable: "Paul's Toyota broke down ..." or "The odometer on Paul's Toyota was approaching 150 000 miles ..." Because people have a natural tendency to explain unexpected events, their completed sentences often imply favorable or unfavorable attitudes toward the targets. Thus, someone who had favorable attitudes toward Toyota might be inclined to try to explain the Toyota's breakdown ("... because Paul drove like a madman."), but not the Toyota's resiliency ("... and he expected it to last for another 150 000 miles."). Someone with negative attitudes toward Toyota might be expected to show the opposite pattern – explaining sentence beginnings featuring good things about Toyota ("... because Paul was the luckiest guy on the planet."), and simply continuing sentence beginnings featuring bad things about Toyota ("... and Paul knew he wouldn't get to the job interview in time.").

As noted above, implicit measures have typically been reserved for use when explicit attitude measures are not expected to perform well, such as when social desirability concerns would influence respondents to report so-called politically correct attitudes, rather than their true feelings. At least one type of paper-and-pencil implicit measure has also been shown to predict unique variance in behavior, beyond what can be predicted by traditional explicit measures – even when the attitude object under consideration is not a socially sensitive topic. Attitudes are complex constructs, with many different facets, each of which may play some unique role in driving behavior. Implicit measures most likely tap different components of attitudes than explicit measures (Gawronski and Bodenhausen, 2006). Traditional, explicit measures are believed to assess propositional components of attitudes – essentially, these are evaluative statements about some object that are stored in memory, such as, "I believe that Apple is a good brand." Response-time-based implicit measures are believed to assess stored evaluations that are automatically activated by the presence of some attitude object. A crude example of an automatically activated concept can be seen in the facility with which "butter" comes to mind when someone says "bread and..." Bread and butter have been paired together so often that invoking bread automatically activates the concept of butter. In the same way that bread and butter are linked, many attitude objects are linked to evaluations (e.g., puppies + good). These evaluations may not be consciously endorsed, but they are automatically activated in the presence of the attitude object, and this activation is reflected in

the speed with which someone might categorize rainbows as good following a cuddly puppy prime in a priming implicit measure.

Research on implicit attitude measures is exceptionally fast-moving. In 1995, there were fewer than 10 published articles on implicit attitudes, but just a decade later there were nearly 100 articles published on the topic. It is an exciting area that continues to develop, and is likely to have an increasingly profound impact on consumer research.

Implicit Processes

As noted above, people cannot possibly consciously process all the advertising and marketing-related stimuli they encounter, much less than all the other stimuli they are faced with. A great deal of cognition goes on outside of our conscious awareness. When it comes to consumer cognition, we can lack conscious awareness of (i) environmental stimuli that may influence our behavior; (ii) automatic cognitive processes that are driving our behavior; and (iii) outcomes, including preferences, judgments, behaviors, emotions, and so forth (Chartrand, 2005). In this section, we review additional work on implicit consumer cognition, beginning with the influence of environmental stimuli that are outside of conscious awareness.

Lacking awareness of stimuli. "I saw a subliminal advertising executive, but only for a second," (attributed to Steven Wright, comedian). One of the greatest obstacles to successfully marketing goods and services is the consumer's resistance to new information and persuasion (*see* CHOICE MODELS) attempts. When consumers are aware of persuasive intent, they are much more likely to try to counterargue, disengage, or otherwise refuse to go along with the would-be persuader. From a marketer's perspective, it would be great to be able to bypass consumers' resistance to persuasion, and just plant ideas to like and buy specific brands directly in the consumers' minds. From a consumer perspective, the prospect of marketers bypassing our ability to resist is terrifying – it would be just a short step from mind control.

For these reasons, marketers and consumers, alike, seem to share a fascination with the idea of bypassing the conscious mind, persuading without allowing any opportunity to resist. This idea ostensibly came to fruition in September 1957, when James Vicary announced that he had used subliminal persuasion to increase popcorn and soda sales. Vicary claimed to have subliminally presented messages ("Eat popcorn" and "Drink Coca-Cola") during a movie. The messages, presented at 1/3000th of a second, were well below the threshold for conscious perception – moviegoers would not have even realized that they had been commanded to eat and drink. And Vicary claimed that this technique increased sales of popcorn and soda in the theater by 18.1 and 57.5%, respectively. The implications of this "experiment" were profoundly troubling. It would be a few years before *The Manchurian Candidate* was popularized (the book was published in 1959; the film debuted in 1962), but by using Vicary's technique one could imagine nudging an unsuspecting public to do any number of things – work harder for less pay, vote for a crooked politician, discriminate or even wage war against some unfortunate group.

The public was nervous. Letters and editorials appeared in the news, Congress and the Federal Communications Commission discussed the ethical implications of subliminal persuasion, and the National Association of Broadcasters banned SUBLIMINAL ADVERTISING. The public disquiet turned out to be unfounded. Attempts to duplicate Vicary's claims were unsuccessful, and in a 1962 interview in *Advertising Age* Vicary admitted that his claim was bogus. He had made the claim in an effort to promote a company in which he was a partner, the Subliminal Projection Company. Despite no solid evidence in support of subliminal persuasion, the concept remained firmly embedded in the public mind.

Several best-selling books promoted the idea of subliminal persuasion, including Vance Packard's (1957) *The Hidden Persuaders*, and Wilson Brian Key's luridly titled, *Subliminal Seduction: Are You Being Sexually Aroused by this Picture*? (1973) and *The Clam-Plate Orgy: And Other Subliminals the Media Use to Manipulate Your Behavior* (1980). Surveys reveal that roughly three-fourths of Americans are aware of subliminal advertising, and of those

who are aware nearly all believe it is actually used, and close to half believe that they may be susceptible to subliminal advertising.

The truth about subliminal persuasion is a little more down-to-earth. For one thing, the term *subliminal* is very often misused in the general media. Steven Wright's clever line about a subliminal advertising executive exemplifies this misuse. The term subliminal evokes the limen, which is the threshold, or point, at which a stimulus evokes a sensation. A subliminally presented stimulus is one that evokes no sensation in the respondent; that is, the respondent has no conscious sense of having seen (or heard, smelled, tasted, touched) the stimulus. Mr. Wright would not have perceived a truly subliminal advertising executive. As well, the limen varies among people, such that one person's subliminal stimulus may be another's supraliminal (consciously perceived) stimulus; the limen also varies within individuals, so that stimuli presented at the same level of intensity may evoke a response at some times, but no response at other times. Objectively operationalizing a subliminal stimulus is therefore quite difficult. Most researchers have opted for a more subjective approach, in which a stimulus is considered subliminal if the respondent cannot notice its presentation (even when they are highly motivated to do so, such as when they are offered cash incentives for correctly identifying the presence of a stimulus).

Subliminal stimuli should be differentiated from transformed stimuli, or unattended stimuli. The former are distorted, blurred, or otherwise degraded so as to be less recognizable; the latter may be clearly recognizable, but are not intended to serve as focal points for attention. An auditory example of transformed stimuli might be music in which different lyrics are audible when played in reverse. An auditory example of unattended stimuli is when different messages are played in left and right headphone channels, while the respondent is instructed to attend to just one channel. Many of the examples used in books by Key and Vicary are transformed, or unattended, rather than strictly subliminal; so most of the claims by Key and Vicary are misinformed, at best, and downright misleading, at worst.

Good, solid evidence for subliminal persuasion is scarce, but it does exist. A review of over 150 mass media articles and over 200 academic papers obtained no solid support for subliminal persuasion (Pratkanis and Aronson, 1992). The articles were seriously flawed in a variety of different ways, making it difficult to infer valid and reliable effects of subliminal stimuli. However, several recent academic articles have offered some evidence that subliminally presented stimuli can have an impact on consumer behavior.

In one study, participants watched an episode of *The Simpsons*. Half of the participants were presented with subliminal primes pertaining to thirst, and the other half were not. Participants who were exposed to the thirst primes later rated themselves as being thirstier than those who were not exposed to the thirst primes (Cooper and Cooper, 2002). In another study, the researchers experimentally manipulated thirst by having all of the participants eat cookies and having half of the participants then drink a glass of water. Next, participants were subliminally exposed to thirst-related words, or words unrelated to thirst. In this study, there was no effect of the thirst prime on participants' self-reported thirst; however, in a fake taste test, participants who were both thirsty and primed with thirst-related words drank the most Kool-Aid (Strahan, Spencer, and Zanna, 2002). In yet another study, researchers experimentally manipulated participants' thirst, and then subliminally exposed participants to a particular brand of beverage or neutral words. Participants who were thirsty and exposed to neutral words expressed no preference for one beverage over another, but those who were thirsty and exposed to the branded beverage showed a marked preference for the brand of beverage that had been subliminally primed (Karremans, Stroebe, and Claus, 2006).

Work on subliminal persuasion continues, but the evidence to date suggests that subliminally presented stimuli most definitely cannot induce moviegoers to get up during a film, go to the concessions stand, and buy a tub of popcorn and a Coca-Cola. Subliminal stimuli seem to be most influential when people are already motivated to engage in some behavior – the subliminal stimuli tend to nudge people further in the direction in which they are already leaning.

Lacking awareness of cognitive processes. Supraliminal stimuli can also influence consumer judgment, emotion, and behavior. A popular classroom demonstration of this effect begins with a list of words that students are instructed to memorize: ocean, waves, gravity, beach, and moon. In order to prevent students from rehearsing the words (the most common strategy for memorizing a handful of items), the instructor asks students a series of questions and instructs them to answer each question aloud. "Name a brand of automobile." In a crowded classroom, the response is an indistinguishable rumble of different brands. "Name a brand of shoes." Again, the response is an indistinguishable rumble. After a few more questions the instructor says, "Name a brand of laundry detergent." The class responds, almost unanimously, "Tide."

When asked about the surprising unanimity of the Tide response students often volunteer that Tide is a brand leader, or that their family uses Tide, or that it just came to mind. Usually, at least one student will also make the connection between the list of words and the conceptually related brand name. The word list serves as a prime, increasing accessibility of the concept of tides. When an opportunity to use the accessible concept arises, many students will do so. In theory, the laundry detergent brand "Cheer" could be elicited by using the words stadium, fan, mascot, game, and team. Priming is an example of a case where people are aware of some stimulus, but generally unaware of the cognitive processes mediating their behavior, or that the prime has had an influence on their behavior (for a thorough review of accessibility effects, see Wyer, 2008).

Priming a familiar brand increases the likelihood that the brand enters the consumers' consideration set and is selected as a final choice (Coates, Butler, and Berry, 2006). A single exposure to a previously unfamiliar brand name is sufficient to induce a reliable increase in the selection rate of the primed brand over already familiar brand names. A similar effect has been demonstrated with auditory stimuli. Hearing a brand name once may be enough to increase a brand's familiarity. Increased feelings of familiarity may provoke consumers to misattribute their familiarity of the brand to an assumption that the brand is popular and well known. If the exposure context is unable to be retrieved, the brand name tends to be considered an existing established brand (Holden and Vanhule, 1999). Even asking consumers about their general purchase intentions (e.g., how likely are you to buy a new automobile?) can increase the likelihood that they make a purchase (Morwitz and Fitzsimons, 2004). This "mere-measurement" effect seems to operate, like more traditional priming effects, by increasing the accessibility of relevant attitudes and beliefs about the object in question.

Priming has been shown to directly influence consumer behavior, as well. Participants primed with prestige goals were more likely to choose relatively expensive and prestigious socks with the Nike brand; participants primed with thrift goals were more likely to choose relatively inexpensive and less prestigious socks with the Hanes brand (Chartrand *et al.*, 2008). Primes do not have to be presented supraliminally in order to have an effect on behavior: participants subliminally primed with the Apple computer logos showed higher creativity scores than those exposed to IBM logos, and participants primed with the Disney channel reported more honest responses to social undesirable questions than did those primed with E! (Entertainment network; Fitzsimons, Chartrand, and Fitzsimons, 2008).

Implicit memory. Marketers have primarily focused their attention on retrieval, or explicit memory (Shapiro and Krishnan, 2001). Explicit memory is assessed via respondents' conscious recollection of a prior exposure to some stimulus, and an intentional attempt to access the information acquired previously. A great deal of consumer behavior is driven by explicit memory (e.g., "Remember to buy eggs and milk," and "I liked the pie at that restaurant."), but implicit memory also has an important role in consumer behavior. *Implicit memory* may be defined as the unintentional, unconscious retrieval of previously acquired information. It is frequently observed by increased performance on indirect tasks that do not demand conscious recollection of past experience. An example of an implicit memory test comparing amnesiac and

non-amnesiac respondents was provided at the beginning of this article.

In assessing ADVERTISING EFFECTIVENESS, Lee (2002) argued that there are two distinct types of implicit memory, perceptual and conceptual, and that these distinct types of implicit memory can be tapped by different measures. She created conceptual primes by having respondents read brand names in sentences (i.e., in a context where conceptual features are salient), and perceptual primes by having respondents read brand names in isolation (i.e., perceptual features of the brand name are salient). Further, consumer behavior was differentially impacted by conceptual and perceptual primes; the former increased the likelihood that consumers would select a primed brand from memory, whereas the latter increased the likelihood that consumers would select a primed brand from a physical layout of different brands. This result has important practical implications for advertisers: when customers are expected to retrieve different product alternatives from memory, persuasive messages should provide relevant contexts to motivate elaborative processing of the brand or product. Otherwise, displaying a prominent product image (perceptual cues) might be a better strategy.

Researchers have supported implicit memory measures as alternatives to explicit recall tasks by showing dissociations between explicit and implicit measures. When participants were asked to do dual tasks (i.e., listen to a short story audio program while advertisements were displayed on slides), their inability to keep their attention focused on the advertisements at the moment of exposure caused different memory performance later in the explicit recall tasks. The number of correct recognitions among participants who were asked immediately after the exposure was quite high (significantly greater than chance levels); however, those who received the recognition task after a week failed at explicit memory retrieval. On the contrary, indirect memory measures were unaffected by either the week delay or attentive states. Overall, implicit tasks revealed an increased likelihood of choosing the advertised brands in all conditions (Shapiro and Krishnan, 2001).

Advertising and product placement seem to have independent and distinct effects on implicit and explicit memory. It is not yet clear whether information that is consciously accessible (explicit memory), and consciously inaccessible (implicit memory) can have additive, or interactive, effects on consumer behavior, but future research should illuminate further effects of implicit memory.

Lacking awareness of behavior. Perhaps one has found oneself unpacking groceries at home, feeling surprised at finding an item come out of his/her grocery bag that he/she does not remember selecting from the shelf: a jar of peanut butter or some Ziploc baggies. Certainly most people have, at one time or another, found themselves parking their own car in their own driveway without any conscious recollection of driving the last few miles home. We have an astounding ability to conduct routine behaviors on autopilot. Riding a bicycle is a remarkably complex task involving the coordinated tasks of steering, balancing, pedaling, breathing, and so on – and it is a task that most people can do without even thinking about it. These are examples of behavior without awareness.

People may be unaware of different aspects of their behavior. Often we are unaware of basic physiological processes, such as breathing. At other times, we are unaware of simple nervous habits, such as repetitive foot tapping. We may also be unaware that a behavior is contingent upon some previous event, such as the effect of a prime. Studying behavior that is outside of awareness is difficult because there are so many ways in which behavior may be outside of awareness – there is no generally agreed upon conceptualization of what, exactly, behavior without awareness actually means.

Still, it is interesting to consider how, and the extent to which, consumer behavior operates outside of conscious awareness. This is an undeveloped area that is wide open for future researchers. As a whole, implicit consumer cognition is a relatively new area of research, so there are numerous gaps where researchers may add substantial contributions. It is an exciting and quickly developing area, with new findings appearing in academic journals monthly.

Bibliography

Chartrand, T. L. (2005) The role of conscious awareness in consumer behavior. *Journal of Consumer Psychology*, **15**, 203–210.

Chartrand, T.L., Huber, J., Shiv, B., and Tanner, R.J. (2008) Nonconscious goals and consumer choice. *Journal of Consumer Research*, **35**, 189–201.

Coates, S., Butler, L., and Berry, D. (2006) Implicit memory and consumer choice: the mediating role of brand familiarity. *Applied Cognitive Psychology*, **20**, 1101–1116.

Cooper, J. and Cooper, G. (2002) Subliminal motivation: a story revisited. *Journal of Applied Social Psychology*, **32**, 2213–2227.

Fitzsimons, G., Chartrand, T., and Fitzsimons, G. (2008) Automatic effects of brand exposure on motivated behavior: how apple makes you "think different". *Journal of Consumer Research*, **35**, 21–35.

Gawronski, B. and Bodenhausen, G.V. (2006) Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological Bulletin*, **132**, 692–731.

Greenwald, A.G., Poehlman, T.A., Uhlmann, E.L., and Banaji, M.R. (2009) Understanding and using the implicit association test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, **97**, 17–41.

Haire, M. (1950) Projective techniques in marketing research. *Journal of Marketing*, **14**, 649–656.

Holden, S.J.S. and Vanhuele, M. (1999) Know the name, forget the exposure: brand familiarity versus memory of exposure context. *Psychology and Marketing*, **16**, 479–496.

Karremans, J.C., Stroebe, W., and Claus, J. (2006) Beyond vicary's fantasies: the impact of subliminal priming and brand choice. *Journal of Experimental Social Psychology*, **42**, 792–798.

Lee, A. (2002) Effects of implicit memory on memory-based versus stimulus-based brand choice. *Journal of Marketing Research*, **39**, 440–454.

Morwitz, V.G. and Fitzsimons, G.J. (2004) The mere-measurement effect: why does measuring intentions change actual behavior? *Journal of Consumer Psychology*, **14**, 64–74.

Payne, B.K., Cheng, C.M., Govorun, O., and Stewart, B.D. (2005) An inkblot for attitudes: affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, **89**, 277–293.

Pratkanis, A.R. and Aronson, E. (1992) *Age of Propaganda: The Everyday Use and Abuse of Persuasion*, W. H. Freeman/Times Books/Henry Holt & Co., New York.

Schacter, D., Chiu, C.-Y., and Ochsner, K. (1993) Implicit memory: a selective review. *Annual Reviews in Neuroscience*, **16**, 159–182.

Shapiro, S. and Krishnan, S. (2001) Memory-based measures for assessing advertising effects: a comparison of explicit and implicit memory effects. *Journal of Advertising*, **30**, 1–13.

Strahan, E.J., Spencer, S.J., and Zanna, M.P. (2002) Subliminal priming and persuasion: striking while the iron is hot. *Journal of Experimental Social Psychology*, **38**, 556–568.

Vargas, P.T., Sekaquaptewa, D., and von Hippel, W. (2007) Armed only with paper and pencil: "Low-Tech" measures of implicit attitudes, in *Implicit Measures of Attitudes* (eds B. Wittenbrink and N. Schwarz),Guilford Press, New York, pp. 103–124.

Warrington, E.K. and Weiskrantz, L. (1970) Amnesia: consolidation or retrieval? *Nature*, **228**, 628–630.

Wittenbrink, B. (2007) Measuring attitudes through priming, in *Implicit Measures of Attitudes* (eds B. Wittenbrink and N. Schwarz), Guilford Press, New York, pp. 17–58.

Wyer, R.S. (2008) The role of knowledge accessibility in cognition and behavior: implications for consumer information processing, in *Handbook of Consumer Psychology* (eds C. Haugtvedt, F. Kardes, and P. Herr), Erlbaum, Mahwah, pp. 31–76.

cross-cultural psychology of consumer behavior

Minkyung Koo and Sharon Shavitt

Introduction

As new global markets emerge, and existing markets become increasingly segmented along ethnic or subcultural lines, the need to market effectively to consumers who have different cultural values has never been more important. Thus, it is no surprise that in the last several years, culture has rapidly emerged as a central focus of research in consumer behavior. This development followed on the heels of extensive social psychological research on culture, which provided a strong theoretical foundation for the consumer-behavior studies that followed.

What is culture? Culture consists of shared knowledge that provides the standards for perceiving, believing, evaluating, feeling, communicating, and acting among those who share a language, a historical period, and a geographic location. As a psychological construct, culture can be studied in multiple ways – across nations, across ethnic groups within nations, across individuals within nations (focusing on cultural orientation), and across situations through the priming of cultural values. The dimensions of individualism versus collectivism, independence versus interdependence, and analysis versus holism have in recent years received significant research attention. This attention has resulted in a great number of studies revealing both antecedents and consequences of the cultural differences between East Asian and North American cultures. As discussed subsequently, regardless of how culture is studied, cultural distinctions have important implications for advertising content, persuasiveness of appeals, consumer motivation, and consumer judgment processes.

Article scope and overview This article reviews major cultural constructs and theoretical implications of cultural differences in consumer information processing, judgments, and choices. Our review is necessarily selective, focusing on findings specific to the consumer domain rather than providing a general review of cultural differences (for an excellent general review, see Wyer, Chiu, and Hong, 2009). It should also be noted that because of space limitations, this article does not cover some major topics in cross-cultural consumer behaviors such as self-regulation and risk-taking, as well as methodological issues such as response styles and biases (Shavitt, Lee, and Torelli, 2009; Shavitt, Torelli, and Wong, 2009).

In this article, the cultural constructs of individualism/collectivism and the independent/interdependent self-construals associated with them are given special attention because extensive research has demonstrated the implications of these distinctions for processes and outcomes relevant to consumer behavior. The most recent refinements to these constructs are briefly reviewed in an attempt to identify additional cultural variables likely to enhance the understanding of cross-cultural consumer behavior. We also review cultural differences in thinking styles as a major emerging cultural distinction and focus on their implications for consumer-behavior research. Finally, we close with a review of cross-cultural differences in advertising content and the persuasiveness of appeals.

Key Constructs and Dimensions of Culture

Individualism versus collectivism. The constructs of *individualism* and *collectivism* represent the most broadly used dimensions of cultural variability for cross-cultural comparison. In individualistic cultures, people value independence from others and subordinate the goals of their in-groups to their own personal goals. In collectivistic cultures, in contrast, individuals value interdependent relationships to others and subordinate their personal goals to those of their in-groups (Hofstede, 1980; Triandis, 1995). The key distinction involves the extent to which one defines the self in relation to others. In individualistic cultural contexts, people tend to have an independent self-construal whereby the self is defined as autonomous and unique. In collectivistic cultural contexts, by contrast, people tend to have an interdependent self-construal whereby

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

the self is seen as inextricably and fundamentally embedded within a larger social network of roles and relationships (Markus and Kitayama, 1991).

National cultures that celebrate the values of independence, such as the United States, Canada, Germany, and Denmark, are typically categorized as individualistic societies in which an independent self-construal is common. In contrast, cultures that nurture the values of fulfilling one's obligations over one's own personal wishes, including most East Asian and Latin American countries, are categorized as collectivistic societies in which an interdependent self-construal is common (Hofstede, 1980; Triandis, 1995).

A large body of research in psychology has demonstrated the many implications of individualism/collectivism and independent/interdependent self-construals for social perception and social behavior (Markus and Kitayama, 1991; Triandis, 1995). These findings indicate consistently that individualists and people with an independent self-construal are oriented toward products and experiences that promote achievement and autonomy, offer personal benefits, and enable expression of one's distinctive qualities. On the other hand, collectivists and people with an interdependent self-construal are oriented toward products and experiences that allow one to avoid negative outcomes, maintain harmony and strong social connections with others, and dutifully fulfill social roles.

Although a given self-construal can be more chronically accessible in a particular culture, cultures generally provide sufficient experiences with independent and interdependent views of the self to allow either type of self-construal to be primed (Oyserman, Coon, and Kemmelmeier, 2002; Oyserman and Lee, 2007). Numerous studies have established that these activated self-views impact judgments in ways that parallel cross-national differences (Shavitt, Lee, and Torelli, 2009), for instance, by activating distinct self goals (Lalwani and Shavitt, 2009). People in general, and especially bicultural people, can readily switch back and forth between independent and interdependent cultural frames in response to their contexts. For instance, Lau-Gesk (2003) found that independent (interdependent) self-construals were temporarily activated when bicultural consumers were exposed to individually focused (interpersonally focused) appeals.

In sum, the distinctions between individualistic and collectivistic societies, and independent and interdependent self-construals, are crucial to the understanding of cross-cultural differences in consumer behavior. The studies to be reviewed here offer extensive evidence that these cultural classifications have fundamental implications for consumption-related outcomes.

Refined individualism versus collectivism. The conceptualizations of individualism and collectivism, and independence/interdependence, have historically been broad and multidimensional, summarizing a host of differences in focus of attention, self-definitions, motivations, emotional connections to in-groups, as well as belief systems and behavioral patterns (Hofstede, 1980; Oyserman, Coon, and Kemmelmeier, 2002). In addition, recent studies have proposed useful refinements to these broader cultural categories (Shavitt, Lee, and Torelli, 2009). These studies suggest that the nature and meaning of individualism and collectivism varies across gender and ethnic lines, as well as across family groupings and institutions. Although the breadth of the individualism-collectivism constructs lends integrative strengths, research indicates that further refinements of these categories can enhance the prediction of consumer behavior.

The horizontal/vertical distinction. With in the individualism-collectivism framework, Triandis *et al.* (Triandis, 1995; Triandis and Gelfand, 1998) have recently introduced a further distinction between societies that are *horizontal* (valuing equality) and those that are *vertical* (emphasizing hierarchy), and a scale to measure these orientations at the individual level. The horizontal/vertical distinction emerges from the observation that American or British individualism differs from, say, Norwegian or Danish individualism in much the same way that Japanese or Korean collectivism differs from the collectivism of the Israeli kibbutz. Specifically, in vertical individualist (VI) societies (e.g., United States and Great Britain), people strive to become distinguished and acquire status via competition (Shavitt, Torelli and Wong,

2009); whereas in horizontal individualist (HI) cultural contexts (e.g., Sweden and Norway), people value uniqueness but are not especially interested in becoming distinguished and achieving high status (Nelson and Shavitt, 2002). In contrast, in vertical collectivistic (VC) societies (e.g., Korea and Japan), people emphasize the subordination of their goals to those of their in-groups, submit to the will of authority, and support competitions between their in-groups and out-groups. Finally, in horizontal collectivist (HC) cultural contexts (e.g., exemplified historically by the Israeli Kibbutz), people see themselves as similar to others, and emphasize shared goals and sociability, but instead of submitting to authority, their view of power focuses on benevolence and helping others (Shavitt, Lee, and Torelli, 2009).

However, the modal comparisons in consumer research are between the United States (VI) and any of a number of Pacific Rim countries (VC). This means that much of what is known about consumer behavior in individualistic and collectivistic societies reflects vertical forms of these syndromes and may not generalize, for example, comparisons between Sweden (HI) and Israel (HC) or other sets of horizontal cultures. As an example, conformity in product choice, as examined by Kim and Markus (1999), may be a tendency specific to VC cultures, in which deference to authority and to in-group wishes is stressed. Much lower levels of conformity may be observed in HC cultures, which emphasize sociability but not deference (Triandis and Gelfand, 1998). Thus, differences in consumers' conformity between Korea (VC) and the United States (VI) may not characterize broad individualism-collectivism differences, because levels of product conformity in HC contexts might not exceed those in HI contexts.

Indeed, several recent studies of this horizontal/vertical cultural distinction have provided evidence for its value as a predictor of new consumer psychology phenomena and as a basis for refining the understanding of known phenomena (Shavitt *et al.*, 2006). For instance, Lalwani, Shavitt, and Johnson (2006) showed that differences in the self-presentational responses observed for individualists and collectivists are mediated at the individual level by the horizontal but not the vertical versions of these cultural orientations. This suggests that culturally linked self-presentational efforts reflect distinct goals of being seen as self-reliant and capable (valued in HI contexts) versus sociable and benevolent (valued in HC contexts).

Further evidence for the value of the horizontal-vertical distinction comes from a study of country-of-origin effects. Gürhan-Canli and Maheswaran (2000) demonstrated that the tendency to favor products from one's own country over foreign products emerged more strongly in Japan (a VC culture) than in the United States (a VI culture). Mediation analyses using individual consumers' self-rated cultural values further indicated that only the vertical aspect of individualism and collectivism accounted for the country-of-origin effects in Japan. In other words, the collectivistic tendency to favor one's own country's products appeared to be driven by cultural values that stress hierarchy, competition, and deference to in-group wishes, not by values that stress interdependence more generally.

In line with this, as noted earlier, research suggests that mental representations of power in terms of status and competition versus benevolence differ reliably between vertical and horizontal cultural backgrounds and orientations. These differences impact consumer information processing and the interpretation of power-related stimuli (Shavitt, Lee, and Torelli, 2009). Finally, content analyses of magazine advertisements in several countries suggested that status-oriented themes of hierarchy, luxury, prestige, and distinction were generally more prevalent in societies presumed to have vertical cultural profiles (e.g., Korea, Russia) than a horizontal cultural profile (Denmark) (Shavitt *et al.*, 2006).

Culture and thinking styles. East Asian and North American cultural differences have been well documented in social psychological research, especially in terms of the differences in individualistic-collectivistic values and independent-interdependent self-systems (Markus and Kitayama, 1991). Many of these cross-cultural studies of consumer behavior have provided evidence that advertising (such

as magazine ads, Internet advertising, and TV commercials) from Western cultures is in general more individualistic and less collectivistic than advertising from Asian cultures (Morling and Lamoreaux, 2008), and that consumers from Western cultures are more likely to be persuaded by individualistic ads and those from East Asian cultures are more likely to be persuaded by collectivistic ads (Han and Shavitt, 1994). However, relatively little research has been done on cross-cultural differences in consumers' thinking orientations. The following section provides a general review of cultural differences in thinking styles in addition to the findings of relevant studies of consumer behavior and advertising effects.

Analytic versus holistic thinking. Broadly speaking, Westerners tend to adopt an analytic thinking style that emphasizes the independence of individual objects, whereas East Asians tend to adopt a holistic view emphasizing that the world is composed of interrelated elements (Nisbett *et al.*, 2001). The analytic style of Westerners and the holistic style of East Asians have been demonstrated in various cognitive domains such as attention, causal reasoning, perception of change, tolerance of contradiction, and categorization.

The analytic style of attention is field independent (mainly oriented toward an object itself), whereas holistic attention is field dependent (focused on the relationship between objects and/or the field in which they are embedded) (Nisbett *et al.*, 2001). This difference in the orientation of attention is also seen in the way East Asians and Westerners perceive and explain social events. East Asians tend to assume that each element in the world is somehow intertwined, and thus an event or object can be understood only in the context of the whole set of relevant factors. By contrast, Westerners tend to explain a certain event in terms of direct causal links, thereby considering fewer reasons than East Asians, who tend to consider a broader set of reasons, regardless of their relevance to the event (Choi *et al.*, 2003).

Furthermore, in explaining causality of a social event, analytic thinkers tend to focus on the internal dispositions of an actor, whereas holistic thinkers tend to consider a broader set of reasons (including both dispositional and contextual information) and are therefore less likely to attribute an outcome to an actor's internal characteristics (Nisbett *et al.*, 2001). This has implications for brand judgments, as well. Monga and John (2007) found that negative publicity influences analytic (vs. holistic) thinkers more heavily, and thus changes their beliefs about a brand to a greater degree because analytic thinkers are less likely to consider contextual information, and thus are more likely to attribute negative product information to the brand.

From the analytic perspective, objects exist independently, and thus the essence of the objects is stable over time. This assumption promotes a linear perception of change in which no drastic deviation is expected in the pattern of stability or change of a phenomenon (Nisbett, 2003). By contrast, the holistic view of the world assumes that objects are interrelated, and therefore it is less likely that a phenomenon will remain stable over time. This perspective results in a cyclic perception in which people tend to predict fluctuating trends for an event. For example, in predicting future stock-market trends and making investment decisions, Canadians are more likely to make judgments based on recent trends than are Chinese people; thus, when compared to the Chinese, Canadians are more willing to buy stocks when they are in an increasing trend and less willing to buy when stock prices are decreasing (Ji, Zhang, and Guo, 2008).

The cyclic perception of change and expectation of instability prevalent among East Asians renders a Yin-Yang belief that a characteristic of an object can potentially transform into its opposite. Consequently, East Asians tend to hold a dialectical perception in which apparently opposing concepts can simultaneously be true and can peacefully coexist (Nisbett *et al.*, 2001). When confronted with opposing propositions, East Asians tend to resolve contradictions by choosing a middle ground, whereas Westerners to rely on formal logic in resolving contradictions by choosing one of the opposing propositions. For example, US consumers tend to resolve incongruities with an attenuation strategy in which one piece of information is favored over another inconsistent piece of information. In contrast, Hong Kong Chinese consumers tend

to follow an additive strategy in which both pieces of information are combined to influence judgments (Aaker and Sengupta, 2000).

East Asians and Westerners also perceive conflicting emotions in different ways. For example, Bagozzi, Wong, and Yi (1999) showed that Chinese tend to hold a dialectical perception that pleasant and unpleasant emotions can be experienced at the same time. Thus, their frequency judgment for pleasant emotions is positively correlated with their frequency judgment for unpleasant emotions. By contrast, this study found that for Americans the perceived frequency of pleasant emotions is inversely correlated with the perceived frequency of unpleasant emotions. Schimmack, Oishi, and Diener (2002) analyzed 38 nationalities and demonstrated that this cultural difference results from dialectical thinking, not from a difference in individualistic-collectivistic values. Moreover, Williams and Aaker (2002) demonstrated that opposing emotions (e.g., both happiness and sadness) in persuasion appeals elicit more positive attitudes among Asian Americans than among European Americans.

Westerners pay more attention to individual objects and attribute causality to them, whereas East Asians focus more on the field. Westerners are more accustomed to formulating rules that govern internal properties of objects and tend to categorize things by applying those rules. In contrast, East Asians organize objects on the basis of their relationship to other objects or to the field (Nisbett, 2003), and therefore they tend to categorize objects according to their overall similarities. Thus, when presented with pictures of a panda, a monkey, and a banana, East Asians tend to categorize the monkey and banana together based on the relationship between the two, whereas Westerners tend to categorize the panda and monkey into one group based on the traits that characterize them (Ji, Zhang, and Nisbett, 2004).

Cultural differences in the way people categorize objects (rule/trait-based vs. similarity/relationship-based) also appear in the way they organize and store brand information. For example, Ng and Houston (2006) showed that Americans are less likely to retrieve brand exemplars (i.e., specific products or subcategories) than brand beliefs (i.e., general descriptive or evaluative thoughts), whereas the reverse was the case for Singaporeans. These results emerged from an analytic tendency to focus on "global beliefs" abstracted from prior product experiences and a holistic tendency to focus on contextual and incidental details about the product. Similarly, Monga and John (2008) found that, compared to Americans, Indians tend to perceive a higher degree of fit between a parent brand (e.g., Kodak) and its brand extension (e.g., Kodak filing cabinet, Kodak greeting cards), and to evaluate the brand extension more positively. This result reflects Indians' holistic tendency to base their judgments more heavily on the relationships between brand extensions and parent brands than do their American counterparts.

A variety of methods and techniques have been developed to measure cultural differences in thinking styles (Choi, Koo, and Choi, 2007; Ji, Zhang, and Nisbett, 2004; Monga and John, 2007), including responses to cognitive tasks, scenarios and questions, physiological measures, a scale, and analyses of various cultural products. Furthermore, priming an independent versus interdependent view of self has also been found to promote analytic and holistic modes of thinking, respectively. For example, people primed with an independent self-view were more likely to focus on a focal object and thus were better at finding an embedded figure by separating the figure from its background than were those primed with an interdependent self-view (see Oyserman and Lee, 2007, for a review).

Additional dimensions. Numerous other cultural distinctions deserve further attention in consumer research. A focus upon these relatively under-researched constructs as antecedents may allow for broadening the range of cultural differences beyond those currently investigated. For instance, Schwartz's (1992) circumplex structure of values, which is highly robust cross-nationally, parallels the HI, VI, HC, VC typology and offers a particularly detailed and comprehensive basis for classification. In his large-scale studies of work values, Hofstede (1980) derived three other dimensions of cultural variation in addition to individualism: *power distance* (acceptance of power inequality in

organizations, a construct conceptually relevant to the vertical/horizontal distinction), *uncertainty avoidance* (the degree of tolerance for ambiguity or uncertainty about the future), and *masculinity/femininity* (preference for achievement and assertiveness versus modesty and nurturing relationships). Indeed, individualism was the second dimension identified by Hofstede (1980), whereas power distance emerged as the first dimension. A few marketing-oriented studies have employed Hofstede's nation-level classifications (Shavitt, Lee, and Torelli, 2009), but more potential remains for identifying consequences for consumer judgments and behaviors. For instance, uncertainty avoidance has been conceptualized as a syndrome related to anxiety, rule orientation, need for security, and deference to experts (Hofstede, 1980). As such, one might speculate that the level of uncertainty avoidance in a culture will predict the tendency for advertisements to use fear appeals or appeals to safety and security, and the tendency for advertisements to employ expert spokespersons. Differences along this cultural dimension may also predict patterns in the diffusion of product innovations, particularly innovations whose purchase entails a degree of risk.

Culture and Persuasive Appeals

Most research on cultural influences on judgment and persuasion has examined the implications of individualism/collectivism or independent/interdependent self-construals. In general, the findings suggest that the prevalence or the persuasiveness of a given type of appeal matches the cultural value orientation of the society (Shavitt, Lee, and Torelli, 2009). For instance, appeals to individuality, personal benefits, and achievement are usually more prevalent and persuasive in individualistic compared to collectivistic cultures, whereas appeals to group benefits, harmony, and conformity are usually more prevalent and persuasive in collectivistic compared to individualistic cultures. Such evidence for "cultural matching" in the nature of appeals has been followed by studies examining the distinct psychological processes driving persuasion across cultures. These studies suggest that culture can affect how people process and interpret product-related information. It can determine the type of information that is weighed more heavily for making judgments (e.g., product attributes vs other consumers' opinions). However, brand and product characteristics can constrain the role of cultural variables in information processing and persuasion, with some brands and products serving as stronger carriers of cultural values (Shavitt, Torelli and Wong, 2009).

Cultural differences in the content of message appeals. Cross-cultural content analyses of advertisements can yield valuable evidence about distinctions in cultural values. For instance, American advertisers and consumer researchers often assume that consumer learning about the brand precedes other marketing effects, such as liking and buying the brand. Thus, advertisements that attempt to teach the consumer about the brand are typical in the United States, although other types of advertisements are also used.

In contrast, as Miracle (1987) suggested, the typical goal of advertisements in Japan appears very different. There, advertisements tend to focus on "making friends" with the audience and showing that the company understands their feelings. The assumption is that consumers will buy once they feel familiar with and have a sense of trust in the company. Because Japan, Korea, and other Pacific Rim countries are collectivist cultures that tend toward implicit and indirect communication practices (Triandis, 1995), Miracle suggested that the mood and tone of commercials in these countries will be particularly important in establishing good feelings about the advertiser. Several studies have supported these notions, showing that advertisements in Japan and Korea, compared to those in the United States, rely more on symbolism, mood, and aesthetics and less on direct approaches such as brand comparisons (Shavitt, Lee, and Torelli, 2009). The ads may be equally informative about the brand across cultures. It is the type of appeal that will vary.

For instance, a content analysis of magazine advertisements revealed that in Korea, compared to the United States, advertisements are more focused on family well-being, interdependence, group goals, and harmony, whereas they are

less focused on self-improvement, ambition, personal goals, independence, and individuality (Han and Shavitt, 1994). However, as one might expect, the nature of the advertised product moderated these effects. Cultural differences emerged strongly only for products that tend to be purchased and used along with other persons (e.g., groceries, cars). Products that do not tend to be shared (e.g., health and beauty aids, clothing) are promoted more in terms of personal, individualistic benefits in both countries.

Paralleling the overall cross-national differences, a content analysis by Kim and Markus (1999) indicated that Korean advertisements, compared to US advertisements, were characterized by more conformity themes (e.g., respect for collective values and beliefs) and fewer uniqueness themes (e.g., rebelling against collective values and beliefs). Website content in Eastern and Western countries also appears to differ in the emphasis on individual versus collective activities (Shavitt, Lee, and Torelli, 2009).

Finally, it is important to note that, in countries experiencing rapid economic growth, advertising content does not necessarily reflect existing cultural values, instead promoting new, aspirational values such as individuality and modernity. For instance, in China, in recent years, westernized ad appeals are increasingly common. Appeals to youth/modernity, individuality/independence, and technology are especially salient in Chinese advertisements that target the younger generation (Zhang and Shavitt, 2003). Similarly, during a period of rapid transition in South Korea's economy (1968–1998), content analysis of advertisements revealed substantial shifts toward individualistic, modernity-oriented appeals (Han and Shavitt, 2005).

Cultural differences in judgment and persuasion. Research suggests that the persuasiveness of appeals may mirror the cultural differences in their prevalence. An experiment by Han and Shavitt (1994) showed that appeals to individualistic values (e.g., "Solo cleans with a softness that you will love") were more persuasive in the United States and appeals to collectivistic values (e.g., "Solo cleans with a softness that your family will love") were more persuasive in Korea. Again, however, this effect was much more evident for products that are shared (laundry detergent, clothes iron) than for those that are not (chewing gum, running shoes).

Zhang and Gelb (1996) found a similar pattern in the persuasiveness of individualistic versus collectivistic appeals in an experiment conducted in the United States and China. Moreover, this effect appeared to be moderated by whether the advertised product is socially visible (camera) versus privately used (toothbrush). Finally, Wang and Mowen (1997) showed in a US sample that individual differences in separateness/connectedness self-schema (i.e., the degree to which one views the self as independent of or interconnected with important others) predicts attitudes toward individualistic versus collectivistic ad appeals for a credit card. Thus, cultural orientation and national culture have implications for the effectiveness of appeals. However, such cultural differences would only be anticipated for those products or uses that are relevant to both personal and group goals.

Cultural differences in persuasion are also revealed in the diagnosticity of certain types of information. For instance, Aaker and Maheswaran (1997) showed that consensus information regarding other consumers' opinions is not treated as a heuristic cue by Hong Kong Chinese (as it is in the United States) but is instead perceived and processed as diagnostic information. Thus, collectivists resolve incongruity in favor of consensus information, not brand attributes. This would be expected in a culture that stresses conformity and responsiveness to others' views. On the other hand, cues whose (low) diagnosticity is not expected to vary cross-culturally (e.g., number of attributes presented) elicit similar heuristic processing in the United States and Hong Kong.

Finally, because cognitive associations with power vary with horizontal and vertical cultural orientations and with ethnicity, as noted earlier, Torelli *et al.*, found differences in the interpretive processes and mindsets triggered when power is salient. Specifically, people whose cultural orientation predisposes a status-oriented view of power activate cognitive processes that facilitate defending their power, such as reasserting control by confirming

prior stereotypes about a brand. In contrast, people whose cultural orientation predisposes a benevolence-oriented view of power activate cognitive processes that facilitate helping others, such as by forming accurate, careful impressions of brands (Shavitt, Torelli and Wong, 2009).

Conclusions

As marketing efforts are increasingly globalized, understanding cross-cultural consumer behavior has become a key focus of consumer research. In recent years, research in consumer behavior has addressed a broadening set of cross-cultural issues and dimensions. Research has provided an enhanced understanding of the relations between culture and self-construal, motivation, thinking style, and consumer persuasion. Research has also begun to address the psychological mechanisms underlying cross-cultural differences in consumer judgments, and the products and contexts for which these differences are most likely to be observed. Understanding cultural differences has become crucial for effective marketing and advertising. In future research, it will be important to further distinguish cultural similarities and differences in consumer judgments, identify within-culture or subgroup differences that parallel between-culture differences, and explore their rich implications in consumer behavior.

Acknowledgments

Preparation of this article was supported by Grant #1R01HD053636-01A1from the National Institutes of Health, Grant #0648539 from the National Science Foundation, and Grant #63842 from the Robert Wood Johnson Foundation to Sharon Shavitt.

Bibliography

Aaker, J.L. and Maheswaran, D. (1997) The effect of cultural orientation on persuasion. *Journal of Consumer Research*, **24** (3), 315–328.

Aaker, J.L. and Sengupta, J. (2000) Addivity versus attenuation: the role of culture in the resolution of information incongruity. *Journal of Consumer Psychology*, **9** (2), 67–82.

Bagozzi, R.P., Wong, N., and Yi, Y. (1999) The role of culture and gender in the relationship between positive and negative affect. *Cognition and Emotion*, **13** (6), 641–672.

Choi, I., Dalal, R., Kim-Prieto, C., and Park, H. (2003) Culture and judgement of causal relevance. *Journal of Personality and Social Psychology*, **84** (1), 46–59.

Choi, I., Koo, M., and Choi, J. (2007) Measuring the analytic vs. the holistic thinking style. *Personality and Social Psychology Bulletin*, **33** (5), 691–705.

Gürhan-Canli, Z. and Maheswaran, D. (2000) Cultural variations in country of origin effects. *Journal of Marketing Research*, **37** (3), 309–317.

Han, S.-P. and Shavitt, S. (1994) Persuasion and culture: advertising appeals in individualistic and collectivistic societies. *Journal of Experimental Social Psychology*, **30** (4), 326.

Han, S. and Shavitt, S. (2005) Westernization of cultural values in Korean advertising: a longitudinal content analysis of magazine ads from 1968-1998, in *Advances in consumer research*, Vol. 32 (eds G. Menon and A.R. Rao), Association for Consumer Research, Provo, UT, pp. 249–250.

Hofstede, G.H. (1980) *Culture's Consequences: International Differences in Work-Related Values*, Sage, Newbury, Park.

Ji, L.J., Zhang, Z., and Guo, T. (2008) To buy or to sell: cultural differences in stock market decisions based on stock price trends. *Journal of Behavioral Decision Making*, **21** (4), 399–413.

Ji, L.J., Zhang, Z., and Nisbett, R.E. (2004) Is it Culture, or is it language? Examination of language effects in cross-cultural research on categorization. *Journal of Personality and Social Psychology*, **87** (1), 57–65.

Kim, H.S. and Markus, H.R. (1999) Deviance or uniqueness, harmony or conformity? A cultural analysis. *Journal of Personality and Social Psychology*, **77** (4), 785–800.

Lalwani, A.K. and Shavitt, S. (2009) The "me" I claim to be: cultural self-construal elicits self-presentational goal pursuit. *Journal of Personality and Social Psychology*, **97** (1), 88–102.

Lalwani, A., Shavitt, S., and Johnson, T.P. (2006) What is the relation between cultural orientation and socially desirable responding? *Journal of Personality and Social Psychology*, **90** (1), 165–178.

Lau-Gesk, L.G. (2003) Activating culture through persuasion appeals: an examination of the bicultural consumer. *Journal of Consumer Psychology*, **13** (3), 301–315.

Markus, H.R. and Kitayama, S. (1991) Culture and the self: implications for cognition, emotion, and motivation. *Psychological Review*, **98** (2), 224–253.

Miracle, G.E. (1987) Feel-Do-Learn: an alternative sequence underlying Japanese consumer response to television commercials, in *Proceedings of the L.A. Conference of the American Academy of Advertising* (ed. F.G. Feasley), The University of South Carolina, Columbia.

Monga, A.B. and John, D.R. (2007) Cultural differences in brand extension evaluation: the influence of analytic versus holistic thinking. *Journal of Consumer Research*, **33** (4), 529–536.

Monga, A.B. and John, D.R. (2008) When does negative brand publicity hurt? The moderating influence of analytic versus holistic thinking. *Journal of Consumer Psychology*, **18** (4), 320–332.

Morling, B. and Lamoreaux, M. (2008) Measuring culture outside the head: a meta-analysis of cultural products. *Personality and Social Psychology Review*, **12** (3), 199–221.

Nelson, M.R. and Shavitt, S. (2002) Horizontal and vertical individualism and achievement values: a multi-method examination of Denmark and the U.S. *Journal of Cross-Cultural Psychology*, **33** (5), 439–458.

Ng, S. and Houston, M.J. (2006) Exemplars or beliefs? The impact of self-view on the nature and relative influence of brand associations. *Journal of Consumer Research*, **32** (4), 519–529.

Nisbett, R.E. (2003) *The Geography of Thought: How Asians and Westerners think Differently..., and Why*, Free Press, New York.

Nisbett, R.E., Peng, K., Choi, I., and Norenzayan, A. (2001) Culture and systems of thought: holistic versus analytic cognition. *Psychological Review*, **108** (2), 291–310.

Oyserman, D., Coon, H.M., and Kemmelmeier, M. (2002) Rethinking individualism and collectivism: evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, **128** (1), 3–72.

Oyserman, D. and Lee, S.W.-S. (2007) Priming 'culture': culture as situated cognition, in *Handbook of Cultural Psychology* (eds S. Kitayama and D. Cohen), Guilford Press, New York, pp. 255–282.

Schwartz, S.H. (1992) Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries, in *Advances in experimental social psychology*, Vol. 25 (ed. M.P. Zanna), Academic Press, San Diego, pp. 1–65.

Shavitt, S., Lalwani, A.K., Zhang, J., and Torelli, C.J. (2006) The horizontal/vertical distinction in cross-cultural consumer research. *Journal of Consumer Psychology*, **16** (4), 325–356.

Shavitt, S., Lee, A., and Torelli, C. (2009) New directions in cross-cultural consumer psychology, in *The Social Psychology of Consumer Behavior, a volume in the series, Frontiers of Social Psychology* (eds M. Wänke, A.W. Kruglanski, and J.P. Forgas) Series Editors, Psychology Press, New York, pp. 227–250.

Shavitt, S., Torelli, C., and Wong, J. (2009), Identity-based motivation in a consumer context. *Journal of Consumer Psychology*, **19** (3), pp. 261–266.

Schimmack, U., Oishi, S., and Diener, E. (2002) Cultural influences on the relation between pleasant emotions and unpleasant emotions: Asian dialectic philosophies or individualism-collectivism? *Cognition and Emotion*, **16** (6), 705–719.

Triandis, H.C. (1995) *Individualism and collectivism*, Westview Press, Boulder.

Triandis, H.C. and Gelfand, M.J. (1998) Converging measurement of horizontal and vertical individualism and collectivism. *Journal of Personality and Social Psychology*, **74** (1), 118–128.

Wang, C.L. and Mowen, J.C. (1997) The separateness-connectedness self-schema: scale development and application to message construction. *Psychology and Marketing*, **14** (2), 185–207.

Williams, P. and Aaker, J. (2002) Can mixed emotions peacefully co-exist? *Journal of Consumer Research*, **28** (4), 636–649.

Wyer, R.S., Chiu, C.-Y., and Hong, Y.-Y. (eds) (2009). *Understanding Culture: Theory, Research and Application*, Psychology Press, New York.

Zhang, Y. and Gelb, B.D. (1996) Matching advertising appeals to culture: the influence of products' use conditions. *Journal of Advertising*, **25** (3), 29–46.

Zhang, J. and Shavitt, S. (2003) Cultural values in advertisements to the Chinese X-generation: promoting modernity and individualism. *Journal of Advertising*, **32** (1), 23–33.

consumer decision making

Haiyang Yang and Ziv Carmon

Introduction

The standard benchmark for models of consumer decision making against which all other models are typically compared is that of *homo economicus* (economic man). This elegant model suggests that consumers are rational actors who evaluate all relevant choice alternatives, assess the utility (value) that each can provide, and choose the one that they expect to provide the most utility. Different versions of this prescriptive model vary in their assumptions and in their descriptive fidelity (the extent to which they accurately describe how consumers actually think and behave), but on the whole, that view of consumer decision making is overly simple. For example, it assumes that consumers are sophisticated thinkers who process information effectively, efficiently, and without bias, and know their preferences.

Decades of descriptive research about consumer judgment and decision making shows that consumers systematically deviate from this model in important ways. Consumers' capacity to process information is limited, yet they commonly face a great deal of information and must therefore trade-off the accuracy of their decisions and the mental effort involved. Rather than choosing the optimal option (the one that they expect will maximize utility), many tend to select an option that is "good enough". Moreover, consumers often do *not* know their preferences. They construct rather than reveal their preferences when those are elicited, and these constructed preferences are sensitive to the decision context and the manner in which these preferences are elicited (Bettman, Luce, and Payne, 1998). In addition, consumers are influenced by meta-goals such as a desire to manage negative emotions like potential regret or conflict, or the extent to which their choice seems easy to justify. Complicating the picture further, rather than simply maximizing pleasure and minimizing pain, consumers sometimes pursue higher order goals such as bolstering their self-esteem, achieving a sense of mastery, or seeking deeper meaning.

In this chapter, we review some of the current knowledge on consumer decision making, highlighting a few ways in which consumers deviate from *homo economicus*. The review is *very* selective as we were restricted in how much we could write and cite, allowing us to cover and to acknowledge few of the papers that contributed to the body of knowledge that we describe. In the following sections we first discuss the impact of decision task and context on consumer preference, and then look at how the interplay of affect and cognition shapes consumer decision making. We also review factors influencing consumers' forecasts of future consumption experiences and discuss the relationship between utility maximization behavior and consumer well-being.

Decision Task and Context

Decisions are easy when consumers know what they want. Unfortunately that is frequently not the case; there are many instances in which consumers have but a sense of what they prefer, and must construct their preferences "on the fly" when a judgment or a choice is called for. Because consumers are affected by a variety of perceptual and judgmental biases, and since they are also often unable or unwilling to invest much time, effort, or attention in constructing their preferences, these preferences can be sensitive to how they were elicited. Examples of significant preference elicitation characteristics discussed below include choice set size, attribute quantity, option similarity and justifiability, presentation format, and response mode.

Choice set size and attribute quantity. Consumers who face a complex judgment or decision frequently use heuristics, decision strategies that simplify the task and are hoped to lead to a decision that is close to the best possible one (see Bettman, Luce, and Payne, 1998). Some heuristics avoid trading off attributes against one another. Such strategies are named *noncompensatory*, as an advantage in one attribute cannot compensate for a disadvantage in others. A lexicographic strategy is a notable example: consumers select the option with the highest value on the attribute that they consider most important, ignoring information about other attributes (in case of ties, they also consider

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

the second most important attribute, etc.). Elimination by aspects is another noncompensatory strategy: consumers screen the choice options, and at each round of screening they select the most important attribute and eliminate options that do not exceed a minimum level on that attribute; this process continues until a single option remains. Compensatory strategies, on the other hand, do trade-off among attributes: a high value on one attribute can thus compensate for a low value on another. An example of a compensatory strategy is weighted adding: consumers evaluate each attribute of an option and compute a weighted total score for the option; after obtaining the scores for all the options in the set, consumers can then identify the option with the highest rating. Note that the examples we describe are "pure" strategies, whereas in reality, consumers may use a mixture of strategies or simplified variants of the strategies for a given decision.

Consumers tend to believe that to reach a good decision it is desirable to examine many alternatives, and seek as much information as they can reasonably find about these choice alternatives. However, information about many attributes or choice sets consisting of many options can make decisions too difficult. To illustrate, Iyengar and Lepper (2000) found that more consumers were drawn to a sampling booth of gourmet jams when it displayed 24 different types of jams, compared to when only 6 types were displayed. However, when only 6 jams were displayed more purchases were made – 10 times more than when 24 jams were displayed. These researchers also found that consumers were more satisfied with their choices when they chose from smaller sets of options. In conclusion, it appears that although people often want more options and more information, these are not necessarily beneficial.

Option similarity and justifiability. Consumer decisions are influenced by the extent to which choice options seem similar to one another. For instance, consider consumers choosing between two shirts: *A* that is comfortable to wear but not very stylish, and *B* that is stylish but not so comfortable. Adding to the choice set another shirt, *C*, that is slightly less comfortable and stylish than *A*, but more comfortable than *B*, can increase the attractiveness and relative choice share of shirt *A*. This phenomenon, known as the *asymmetric dominance effect* (attraction effect), violates the economic principle of regularity: adding an alternative to a choice set should not increase the choice share of any option in the original set (Huber, Payne, and Puto, 1982). This effect occurs because presence of the dominated option (shirt *C*) increases the appeal of the dominating option (shirt *A* in our example) as contrast between the former and the latter options helps the trade-off between the different dimensions (style and comfort) appear more favorable to the dominating option (shirt *A*) versus the alternative (shirt *B*).

Consumers are influenced by the extent to which they feel that they can easily justify their decisions, a notion named *reason-based choice*. This notion predicts a variety of interesting phenomena, such as a preference for so-called compromise options (Simonson, 1989). To illustrate, consider our shirt example; if we add to the original choice set of shirts *A* and *B* a shirt *D* that is more comfortable but less stylish than *A* and *B*, both the attractiveness and choice probability of shirt *A* are likely to increase. This is because consumers tend to find choosing an intermediate option (shirt *A*) easier to justify than extreme options (shirt *B* and *D*). This phenomenon is known as the *compromise effect* or *extremeness aversion*.

Presentation format. Consumer decision making is often influenced by how information is organized and presented, as consumers tend to focus on the information that is explicitly displayed and use it in the form that it is displayed. For example, consumers can use price information better when product and price information is presented in a list sorted by unit price than when the same information is displayed in a traditional store setting, price tags on product shelves (Russo, 1977). That said, consumers do sometimes restructure information to make it more suitable for decision making when it seems too arduous to follow the given format.

Information can be displayed so as to establish different standards of comparison (referred to as *reference points*), which, in turn, influence consumer decision making. For example, if consumers are offered a full-featured product

and asked to delete features that they do not want, they tend to choose significantly more features than if they are given a base model and asked to add the features they desire (Park, Jun, and MacInnis, 2000). This tendency has been attributed to two important notions in consumer decision making: reference dependence and loss aversion. According to reference dependence, consumers evaluate options by comparing them to reference points rather than evaluating them in absolute terms. In our example, the configuration that consumers first see serves as the reference point against which subsequent configurations are compared. The notion of loss aversion suggests that giving up an item, a feature, or a level of a feature, hurts more than acquiring the same thing feels good; in other words, losing something looms larger than gaining the same thing (Kahneman and Tversky, 1979). In our example, loss aversion drives consumers to add more features than they detract.

Loss aversion can also lead to a preference for the status quo. For instance, in countries where organ donation is the default decision (status quo) and people are given the option to opt out, consent rates tend to be very high, while in countries that generally seem comparable where the default decision is not to donate (an alternative status quo) and people can choose to opt in, consent rates are much lower (3–20 times lower according to Johnson and Goldstein, 2003).

Even arbitrary reference points can impact decision making. For example, Tversky and Kahneman (1974) used a roulette wheel to randomly generate a number, then asked participants in their experiment whether the number of African countries in the UN was higher or lower than that number, and then asked the participants to estimate the actual number of African UN member countries. Those who first saw a large (roulette derived) number estimated the number of African countries in the UN as significantly greater than those who saw a small (roulette derived) number.

More recent research showed that arbitrary anchors can also influence assessments of a sequence of items. The subsequent valuations tend to be "coherent" with respect to perceived differences among the products, "the entire pattern of valuations can easily create an illusion of order, as if it is being generated by stable underlying preferences" (Ariely, Loewenstein, and Prelec, 2003, p. 73). For example, the willingness to pay for a single bottle of wine was shown to be heavily influenced by an arbitrary number, but bids for multiple units of the same wine followed a reasonable logic: if willingness to pay for one bottle was $X, the willingness to pay for two bottles was a bit less than 200% of $X, for three bottles it was a bit less than 150% of the bid for two bottles, and so on. This phenomenon was named *coherent arbitrariness*, as the first bid is arbitrary but subsequent bids follow a coherent pattern.

Consumers' judgment of an option can also be shaped by whether the option is presented and evaluated separately or jointly with other options; if a choice option is rated favorably in singular evaluation, consumers may prefer the option less if the option is evaluated jointly with other options; conversely, if the option is not rated very positively in singular presentation and evaluation, its rating may improve in joint judgment (Hsee and Leclerc, 1998).

Meta-cognitive experiences related to presentation format can also influence choice. As an example, when choice information is presented in a fuzzy, difficult-to-read font, the resulting perceptual difficulty (of reading) can be attributed to the choice set. Consumers may infer that choosing between the options is difficult, and this can increase the likelihood of choosing to put off a decision (choice deferral) (Novemsky *et al.*, 2007).

Response mode. Procedure invariance, another important principle in the classical economic decision-making paradigm, requires that equivalent preference elicitation procedures elicit the same preferences. A well-known violation of procedure invariance is the prominence effect (Tversky, Sattath, and Slovic, 1988): prominent (important) attributes are weighed more heavily when people choose among options than when they evaluate the same options in other ways, such as matching. Price matching, for example, asks for the price level at which the options will be equally attractive. To illustrate, when choosing between two programs to reduce traffic accident fatalities, one saving more lives but costing more than the other, people weigh saved lives (the more prominent attribute) more

heavily than cost. They therefore tend to choose expensive programs saving more lives, reflecting high willingness to pay to save a life. But when asked a price matching question, such as at what price the program saving more lives is equivalent to the inexpensive program saving fewer lives (the price beyond which the less expensive program saving fewer lives is preferred), they indicate a price roughly proportional in cost per life saved, reflecting lower willingness to pay to save a life.

As another example, consumers' value assessment of the same item can vary greatly depending on how it is assessed. For instance, the highest sum consumers will pay to obtain an item is typically very different from the lowest sum for which they will part from the item. This is true even if potential differences between buyers and sellers (such as liquidity constraints and information asymmetries) are controlled for, and although both types of value assessments should reflect the value that the consumer associates with owning the item. This robust phenomenon (see Maddux et al., in press, for discussion on variation across cultures), named the *endowment effect* (Kahneman, Knetsch, and Thaler, 1990), is due to a combination of factors. One is loss aversion. Another is that buyers and sellers use different heuristics to assess value, each focusing on and thus assessing different aspects of the potential transaction. Buyers focus on the money they stand to forgo should the transaction take place, and thus consider such things as opportunity costs (e.g., other things they could purchase with the money). Sellers, on the other hand, focus on the item that they would forgo should the transaction take place (Carmon and Ariely, 2000).

As yet another example, hedonic dimensions are weighed less heavily in acquisition decisions (which of several options to obtain) than in forfeiture decisions (which of several options to forgo). To illustrate, consumers wishing to buy a chocolate bar and an equally priced notepad, but who can only afford one, are more likely to buy the chocolate when choosing which option to forgo (e.g., if they have both items in their shopping cart when they notice the budget constraint) than when choosing which to select (e.g., if they have neither item in the cart when they notice the budget constraint; Dhar and Wertenbroch, 2000).

Affect in Decision Making

Affect can have considerable impact on consumer decision making. Choices can be so affect-laden that trade-offs among them evoke strong unpleasant emotions. Some consumer decisions reflect reluctance to experience such emotions. For example, consumers sometimes choose to maintain a status quo so as to avoid the unsettling emotions involved with agonizing over which option they should choose instead (Luce, 1998).

Affective state. Affect can significantly influence decision-making processes. Positive affect can sometimes lead to more efficient decision making, for example. Also, when in a positive mood, consumers can decide more quickly and search less for redundant information. When experiencing negative emotions, on the other hand, consumers may seek to make accurate choices to improve how they feel, and may thus engage in more systematic, effortful information processing (Isen, 2001).

How consumers feel as they decide can also color their evaluations. For example, when assessing an object, consumers partly infer their assessment from how the product experience "seems to feel." This is referred to as the "how-do-I-feel-about-it heuristic," or "affect-as-information" (Pham, 1998; Schwarz and Clore, 1988). Incidental emotions (having little to do with the evaluated product, such as gloominess on a cloudy day), can affect how consumers feel at that point in time and can thus contaminate their product assessments, as consumers partly infer their evaluations from how they feel when they assess the product. Thus, positive incidental emotions can lead to more favorable assessments and the opposite is true of negative incidental emotions. Such effects of incidental emotions can be reduced if their true source is made salient. For example, negative effects of gloominess during a cloudy day on product assessment can be reduced if consumers are first explicitly asked what the weather is like then.

Anticipated regret. Anticipated emotions can also influence consumer decision making. For example, when consumers try to predict how they would feel if they were to have made a poor decision, they are more likely to purchase a product that is currently on sale rather than wait for a better sale; they are also more likely to choose higher priced well-known brands (Simonson, 1992). Bar-Hillel and Neter 1996 describe another interesting example of the impact of anticipated affect: participants were first given a lottery ticket and then asked if they would exchange the ticket for another that had objectively better odds. Surprisingly, most participants did not want to trade due to the regret they expected to feel if their original ticket won. More generally, consumers' concerns about the regret that they might feel should their choice turn out badly drive them to seek "safe" choices (*see* EMOTION).

Self-Control

Consumers regularly face choice situations involving indulgent and utilitarian options. They often choose the short lived satisfaction that comes from indulging themselves at the expense of their longer term goals and interests, even when the negative consequences of their decisions are clear. Owing to the so-called "hot-cold empathy gap," consumers often fail to predict such transgressions when they are in a "cold state" – removed from the tempting situation, in physical distance, in time, or in their thoughts. Being in such a cold state, it is easy to mispredict one's ability to act virtuously in the face of temptation. Worse yet, lapses are often exacerbated by visceral states such as feeling significant deprivation, hunger, or lust (Loewenstein, 1996). For example, consumers who shop for food while they are hungry, tend to buy more than they would otherwise purchase, or will want to consume.

Wertenbroch (1998) distinguishes between vice and virtue goods. Vices (such as rich chocolate cake, or cigarettes) are typically preferred to virtues (such as fresh broccoli, or fat free versions of fatty foods) when consumers focus on the immediate consequences of consumption. To solve their self-control problems, consumers sometimes proactively ration their decisions. For example, they sometimes ration purchase quantity, limiting their stock of vice goods and thus consumption opportunities. Consumers of vice goods seeking to regulate their behavior may thus agree to pay higher unit price (for instance, avoid buying cigarette cartons despite the expected savings). Interestingly, some consumers become so future-oriented that they find it difficult to indulge (Kivetz and Simonson, 2002); to cope with this difficulty, they sometimes precommit to indulge (for example, choose luxury items rather than cash of equal or greater value, as a reward from a loyalty program).

Consumer decision making is shaped by spontaneous affect as well as by cognitive processes. When consumers' cognitive resources are constrained, they are more likely to choose options with higher immediate affective rewards (e.g., chocolate cake); conversely, when processing resources are less constrained, consumers are more likely to choose options with more favorable cognitions (e.g., fruit salad; Shiv and Fedorikhin, 1999). If options are presented vividly (e.g., real chocolate cake and fruit salad are physically placed in front of the consumer), even consumers with sufficient cognitive resources to elaborate on their decisions are likely to be swayed by affective considerations and choose to indulge (*see* IMPULSIVE AND COMPULSIVE BUYING).

Misforecasting Consumption Experiences

Another reason consumers often fail to select those options that seem best for them is that they often mispredict their satisfaction with their choices. In this section, we discuss examples of prediction errors that can lead to suboptimal decisions.

Consumers' decisions often rely on intuitive predictions about future consumption experiences. Forecasts of how satisfied consumers will be with the consequences of their choices can unfortunately be quite poor. To illustrate this, Kahneman and Snell (1992) asked participants in an experiment to consume a serving of plain yogurt every day for eight days. The participants incorrectly predicted increasing dislike of yogurt, whereas most came to like it more (or dislike it less). Note that this experience of

eating yogurt is quite familiar, suggesting that consumers may fair more poorly when it comes to experiences with which they are less familiar.

Another reason that consumers tend to mispredict which decisions will be most satisfying for them has to do with a distinction between how a consumption episode is experienced as it happens and how it is summarized and remembered in retrospect. When consumers retrospectively summarize a consumption episode, they tend to focus only on a few key characteristics, such as the rate at which the experience became more or less pleasant over time, the intensity of the most extreme sensation, and the sensation at the final moment of the experience (Ariely and Carmon, 2000). Summary evaluations tend to represent only few aspects of the actual experience, and can thus be poor representations of the actual episode. Nevertheless, they are the best predictor of future decisions.

Incorrect but self-fulfilling expectations. Consumers' assessment of consumption experiences are influenced by their expectations, as the latter serve as a standard against which the actual experience is compared. Some expectations are based on widely held but incorrect beliefs, such as the notion that lower prices reflect lower quality (in this example the reality is that the empirical correlation between price and quality is negligible). Interestingly, in some instances, expectations that are incorrect can nevertheless be self-fulfilling.

For example, consumers who paid a discounted price for an energy drink that was said to increase mental acuity, derived less benefit from consuming the product (they were able to solve fewer puzzles) versus those who purchased the same product at its regular price (Shiv, Carmon, and Ariely, 2005). Similarly, Lee, Frederick, and Ariely (2006) asked patrons of a pub to evaluate the "MIT Brew" which was made of regular beer to which a few drops of balsamic vinegar was added. When the ingredients were revealed before participants sampled the brew, few liked it. However, when participants received the information immediately after they had sampled the brew, they rated the beer highly and were more likely to choose the special brew over regular beer. In other words, only when they (incorrectly) expected the beer to taste badly did it in fact taste that way.

Impact bias. Impact bias occurs when consumers "overestimate the intensity and duration of their emotional reactions to future events" (Wilson and Gilbert, 2005, p. 131). Focalism, the tendency to fixate on certain attributes or events but ignore (or under weigh) others, is one of the causes of the impact bias. For example, Schkade and Kahneman (1998) found that although people's self-reported life satisfaction was similar whether they lived in the Midwest or in California, when asked to rate a similar person living at the other location, they thought Californians would be more satisfied than Midwesterners. This was explained by people's assessments focusing too heavily on the salient qualities of California (climate and cultural opportunities), and largely ignoring other important determinants of happiness. As another example, consumers tend to recall atypical instances (e.g., the worst experience at a restaurant) and may rely on those rare instances in forecasting how they will react to future consumption events.

A related cause of impact bias is a tendency to neglect or underweigh adaptation to changes in circumstances. In a provocative example, Brickman, Coates, and Janoff-Bulman (1978) describe data suggesting that lottery winners are no happier than the average person, and people who became paralyzed in an accident are no less happy than the average person, a year after the dramatic change to their lives. While this example is controversial and may well overstate people's ability to adapt, due to difficulties in measuring and comparing happiness across people, the general point that, people tend to underestimate their ability to adapt, is widely accepted.

Projection bias. Consumer decisions are also influenced by a projection bias, the inability to sufficiently consider how their preferences may change when circumstances become different (Gilbert, Gill, and Wilson, 2002). For example, when in a "hot" state such as hunger or thirst, consumers tend to overweigh the impact of temporary visceral factors on projections of needs and wants (Loewenstein, 1996). For

example, when sexually aroused, men are significantly more willing to engage in unsafe and morally questionable sexual behaviors than they would otherwise consider. (Ariely and Loewenstein, 2006).

Ironic Effects of Trying to Choose Well

Consumers' attempt to choose the very best alternative has interesting consequences for their satisfaction. Studies by Wilson *et al*. (1993) suggest that careful consideration of choices can cause consumers to pay more attention to less important criteria, and therefore choose less satisfying options. Ironically, maximization (seeking the best option) can not only result in normatively better decision outcomes (e.g., better paying jobs) but also in more negative subjective evaluations of the outcomes (e.g., lower job satisfaction). The lower satisfaction of maximizers is said to be due to heavier reliance on criteria that seem objectively important rather than criteria that feel important to them, and greater fixation on unrealized options (Iyengar, Wells, and Schwartz, 2006). Furthermore, those who maximize tend to be more affected by upward social comparisons (naturally dissatisfying comparisons to others who are in a better position), and experience more depression and regret.

Sometimes consumers experience buyers' regret, unsettling displeasure with having made a choice. Ironically, this is common with choices over which consumers extensively elaborated (made an effort to determine the very best option). This happens because elaboration can induce feelings of attachment to the choice options, a sense of prefactual possession of those options. When consumers eventually choose, they effectively lose the prefactual ownership of nonchosen options, which evokes negative feelings associated with loss and increases the appeal of forgone options compared to their appeal before the choice was made (Carmon, Wertenbroch, and Zeelenberg, 2003).

Similarly, over time, satisfaction with the chosen outcome can decrease for those consumers who considered alternatives before they chose. Ironically, satisfaction with a randomly assigned option, as well as satisfaction of consumers who did not consider alternatives, tends to remain comparatively stable (Ritov, 2006). On a related note, when consumers make hyperopic choices (choices emphasizing long- rather than short-term benefits) of virtue over vice, they may experience increasing regret over time because time attenuates emotions of indulgence guilt, but accentuates wistful feelings of missing out on hedonic pleasures, leading to intensified regret (Kivetz and Keinan, 2006) (*see* CONSUMER WELL-BEING).

Summary

Many thoughtful researchers still defend the view that consumer decisions should be considered rational. In contrast, there is a large and growing body of evidence suggesting that consumers are not the *homo economicus* that the classical economic paradigm portrays. Consumers are limited by finite mental capacity; in many situations they have only a sense of what they prefer and must therefore construct their preferences on the fly, and those preferences tend to be contingent on the characteristics of the decision task and context. Consumers' decisions are often swayed by the affect they experience and anticipate and by their predictions regarding future consumption experiences that are often inaccurate. Consumers' attempt to seek utility-maximizing outcomes, ironically, may harm rather than improve their well-being. Although these insights do not yet make up a coherent theory of consumer decision making, they do suggest that *Homo sapiens*' rationality is predictably fallible.

Consumers' decision biases and errors, such as those we review here, can lead to suboptimal decisions. That said, consumers' decisions are often "good enough." In part this is because the heuristics that consumers use are often reasonably effective, and partly since many decisions, including ones in which consumers invest much energy are largely inconsequential, either because the decision is between good alternatives or because the decision is of little real importance. However, in some situations errors can have serious consequences (such as failing to save enough for retirement). The notion of Libertarian Paternalism (Thaler and Sunstein, 2008) proposes that " . . . knowledge of how people

think be used to design choice environments that make it easier for people to choose what is best for themselves, their families and their society." For example, a simple change to employees' default decision so that they regularly contribute money to their retirement savings account unless they choose to opt out (instead of the default being not to save unless they choose to opt in) could greatly help combat failure to save enough for retirement. The paternalistic flavor is controversial, though its proponents point out that it does *not* affect the freedom to choose.

In conclusion, the debate over whether and to what extent consumer decisions can be considered rational has spurred a large body of research yielding many insights into how consumers decide. Beyond that, however, approaches such as that of Libertarian Paternalism, using research on customer decision making to help consumers, seem more productive. We hope that future research will not only work toward developing a complete theory of consumer decision making but also shed more light on how this knowledge can be used to help consumers become better decision makers.

Acknowledgment

We thank Jim Bettman and Itamar Simonson for helpful comments and suggestions.

Bibliography

Ariely, D. and Carmon, Z. (2000) Gestalt characteristics of experiences: the defining features of summarized events. *Journal of Behavioral Decision Making*, **13**, 191–201.

Ariely, D. and Loewenstein, G. (2006) The heat of the moment: the effect of sexual arousal on sexual decision making. *Journal of Behavioral Decision Making*, **19**, 87–98.

Ariely, D., Loewenstein, G., and Prelec, D. (2003) Coherent arbitrariness: stable demand curves without stable preferences. *Quarterly Journal of Economics*, **118**, 73–105.

Bar-Hillel, M. and Neter, E. (1996) Why are people reluctant to trade lottery tickets? *Journal of Personality and Social Psychology*, **70**, 17–27.

Bettman, J.R., Luce, M.F., and Payne, J.W. (1998) Constructive consumer choice processes. *Journal of Consumer Research*, **25**, 187–217.

Brickman, P., Coates, D., and Janoff-Bulman, R. (1978) Lottery winners and accident victims: is happiness relative? *Journal of Personality and Social Psychology*, **36**, 917–927.

Carmon, Z. and Ariely, D. (2000) Focusing on the forgone: why value can appear so different to buyers and sellers. *Journal of Consumer Research*, **27**, 360–370.

Carmon, Z., Wertenbroch, K., and Zeelenberg, M. (2003) Option attachment: when deliberating makes choosing feel like losing. *Journal of Consumer Research*, **30**, 15–29.

Dhar, R. and Wertenbroch, K. (2000) Consumer choice between hedonic and utilitarian goods. *Journal of Marketing Research*, **37**, 60–71.

Gilbert, D.T., Gill, M., and Wilson, T.D. (2002) The future is now: temporal correction in affective forecasting. *Organizational Behavior and Human Decision Processes*, **88**, 430–444.

Hsee, C.K. and Leclerc, F. (1998) Will products look more attractive when evaluated jointly or when evaluated separately? *Journal of Consumer Research*, **25**, 175–186.

Huber, J., Payne, J.W., and Puto, C.P. (1982) Adding asymmetrically dominated alternatives: violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, **9**, 90–98.

Isen, A.M. (2001) An influence of positive affect on decision making in complex situations: theoretical issues with practical implications. *Journal of Consumer Psychology*, **11** (2), 75–85.

Iyengar, S.S. and Lepper, M.R. (2000) When choice is demotivating: can one desire too much of a good thing? *Journal of Personality and Social Psychology*, **79**, 995–1006.

Iyengar, S., Wells, R., and Schwartz, B. (2006) Doing better but feeling worse: looking for the "best" job undermines satisfaction. *Psychological Science*, **17**, 143–150.

Johnson, E.J. and Goldstein, D.G. (2003) Do defaults save lives? *Science*, **302**, 1338–1339.

Kahneman, D., Knetsch, J.L., and Thaler, R.H. (1990) Experimental tests of the endowment effect and the Coase Theorem. *Journal of Political Economy*, **98**, 1325–1348.

Kahneman, D. and Snell, J. (1992) Predicting a changing taste: do people know what they will like? *Journal of Behavioral Decision Making*, **5**, 187–200.

Kahneman, D. and Tversky, A. (1979) Prospect theory: an analysis of decisions under risk. *Econometrica*, **47**, 313–327.

Kivetz, R. and Keinan, A. (2006) Repenting hyperopia: an analysis of self-control regrets. *Journal of Consumer Research*, **33**, 273–282.

Kivetz, R. and Simonson, I. (2002) Self control for the righteous: toward a theory of precommitment to indulgence. *Journal of Consumer Research*, **29**, 199–217.

Lee, L., Frederick, S., and Ariely, D. (2006) Try it, you'll like it: the influence of expectation, consumption, and revelation on preferences for beer. *Psychological Science*, **17**, 1054–1058.

Loewenstein, G. (1996) Out of control: visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, **65** (3), 272–292.

Luce, M.F. (1998) Choosing to avoid: coping with negatively emotion-laden consumer decisions. *Journal of Consumer Research*, **24**, 409–433.

Maddux, W.W., Yang, H., Falk, C. *et al*. (in press) For whom is parting with possessions more painful? Cultural differences in the endowment effect. *Psychological Science*.

Novemsky, N., Dhar, R., Schwarz, N., and Simonson, I. (2007) Preference fluency in choice. *Journal of Marketing Research*, **44**, 347–356.

Park, C.W., Jun, S.Y., and MacInnis, D.J. (2000) Choosing what I want versus eliminating what I don't want: the effects of additive versus subtractive product option framing on consumer decision making. *Journal of Marketing Research*, **37**, 187–202.

Pham, M.T. (1998) Representativeness, relevance, and the use of feelings in decision making. *Journal of Consumer Research*, **25**, 144–159.

Ritov, I. (2006) The effect of time on pleasure with chosen outcomes. *Journal of Behavioral Decision Making*, **19**, 177–190.

Russo, J.E. (1977) The value of unit price information. *Journal of Marketing Research*, **14**, 193–201.

Schkade, D. and Kahneman, D. (1998) Does living in California make people happy? A focusing illusion in judgments of life satisfaction. *Psychological Science*, **9**, 340–346.

Schwarz, N. and Clore, G.L. (1988) How do I feel about it? Informative functions of affective states, in *Affect, Cognition, and Social Behavior* (eds K.Fiedler and J. Forgas), Hogrefe International, Toronto, pp. 44–62.

Shiv, B., Carmon, Z., and Ariely, D. (2005) Placebo effects of marketing actions: consumers may get what they pay for. *Journal of Marketing Research*, **42**, 383–393.

Shiv, B. and Fedorikhin, A. (1999) Heart and mind in conflict: interplay of affect and cognition in consumer decision making. *Journal of Consumer Research*, **26**, 278–282.

Simonson, I. (1989) Choice based on reasons: the case of attraction and compromise effects. *Journal of Consumer Research*, **16**, 158–174.

Simonson, I. (1992) The influence of anticipating regret and responsibility on purchase decisions. *Journal of Consumer Research*, **19**, 105–118.

Thaler, R.H. and Sunstein, C.R. (2008) *Nudge: Improving Decisions about Health, Wealth, and Happiness*, Yale University Press, New Haven .

Tversky, A. and Kahneman, D. (1974) Judgment under uncertainty: heuristics and biases. *Science*, **185**, 1124–1130.

Tversky, A., Sattath, S., and Slovic, P. (1988) Contingent weighting in judgment and choice. *Psychological Review*, **95**, 371–384.

Wertenbroch, K. (1998) Consumption self-control by rationing purchase quantities of virtue and vice. *Marketing Science*, **17**, 317–337.

Wilson, T.D. and Gilbert, D.T. (2005) Affective forecasting: knowing what to want. *Current Directions in Psychological Science*, **14**, 131–134.

Wilson, T.D., Lisle, D.J., Schooler, J.W. *et al.* (1993) Introspecting about reasons can reduce post-choice satisfaction. *Personality and Social Psychology Bulletin*, **19**, 331–339.

motivation and goals

Hans Baumgartner and Rik Pieters

Motivation deals with the question of why consumers do the things they do. It is frequently contrasted with cognition and affect. While cognition refers to the thinking aspect of functioning, and affect to the feeling aspect, motivation deals with the wanting or striving aspect. Two issues are involved: what people want (the direction of motivation) and how much they want something (the force of motivation). Psychological constructs such as needs, motives, desires, wishes, values, drives, and goals all fall within the purview of motivation. It is frequently difficult to distinguish between these constructs precisely, but in general needs and motives are higher level reasons for engaging in a behavior, usually thought to be limited in number so that classifications of human needs and motives can be developed. Desires and wishes tend to refer to lower level wants, often ones whose feasibility may be questionable, with desires being stronger than wishes. Values are high-level ideals that are considered socially desirable by some segments of society, and they may be used as criteria for evaluating objects, events, and behaviors. Drives are hypothetical internal states (often arising from a discrepancy between a desired and actual state and considered to be aversive) that impel people to action and energize behavior. Finally, goals are internal representations of desirable states that people try to attain (approach goals) and undesirable states that they try to avoid (avoidance goals).

In this article, we focus on goals as the central construct of motivation and use the notion of goal in a general sense ranging from relatively concrete aims that provide specific direction for behavior to more abstract purposes (including motives, needs, and values) that imbue subordinate goals with affect. We emphasize the directional aspect of motivation, since involvement (which captures the intensity aspect) is covered in a separate chapter. We start the presentation with a discussion of goal structure and then move to the dynamics of goal pursuit, including both conscious and nonconscious goal setting and goal striving.

The Structure of Goals

Structural approaches to the study of motivation are well established (see Austin and Vancouver, 1996; Carver and Scheier, 1998; Heckhausen, 1991). These frameworks describe different levels of motivations that people are assumed to have, either in general or in a specific situation, ranging from concrete to abstract motivations, and the linkages between motivations at the same or different levels in the hierarchy in terms of their associative, temporal, conditional, or causal relationships. We begin our discussion with two approaches that are well known but usually not mentioned under the rubric of goals (Maslow's hierarchy of needs and Schwartz's classification of values), although they are closely related since needs and values can be conceived as high-level goals. We then describe means-end chain analysis and its extension to modeling consumer goal structures. We end with a brief discussion of purchase motives.

Needs and values. In general, needs and values are abstract motivational constructs that are linked to the more concrete goals that consumers have in specific life-domains and situations.

Maslow's hierarchy of needs. Maslow (1943) proposed a famous theory of motivation that posits five basic human needs: physiological needs (hunger, thirst, sex), safety needs, love needs (belonging, affection), esteem needs (including both self-esteem based on achievement or independence and esteem from others based on prestige or appreciation), and need for self-actualization. The five needs are hypothesized to be organized into a "hierarchy of relative prepotency" such that lower order needs have to be satisfied before higher order needs can emerge. Applied to marketing, and as described in consumer behavior textbooks, the theory suggests that products are differentially relevant to the five needs (e.g., insurance presumably satisfies the need for safety) and that when appealing to consumers' needs, account has to be taken of where in the need hierarchy a target segment is situated. Among the criticisms of the theory are that the posited hierarchy of prepotency may not be valid and that the theory may not be cross-culturally applicable. Importantly, interest in the basic

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

needs proposed by Maslow has been growing in recent years, with specific theories and measures being developed for each of them, such as the need to belong (Baumeister and Leary, 1995). In the domain of consumer behavior, Belk, Ger, and Askegaard (2003) provide a rich framework of needs and wants and their connections to consumer passions.

The structure of values. Particularly for purposes of market segmentation, research on values has a long tradition in marketing. Market segmentation aims to identify groups of consumers with similar within-group and different between-group "preferences." Values are considered such abstract "preferences" that give direction to consumer behavior in a wide range of situations. The most sophisticated framework is Schwartz's (1992) work on universals in the content and structure of values. Extending Rokeach's (1973) structural theory of instrumental (lower level) and terminal (higher level) values, Schwartz distinguishes 10 different value types, corresponding to basic human motivations. The value types can be arranged in a circular order to reflect their compatibilities and conflicts. Furthermore, the 10 value types can be grouped into four higher order value domains (roughly four quadrants of a circle): self-enhancement (achievement, power) versus self-transcendence (universalism, benevolence), and conservation (tradition, conformity, and security) versus openness to change (self-direction, stimulation), with hedonism falling between the openness to change and self-enhancement values. Schwartz also constructed an instrument in which participants rate over 50 individual values based on importance, and research using this instrument, involving cultures from around the globe, suggests that the hypothesized structure of values may be universal.

Hierarchical conceptualizations of goals.

Means-end chain analysis. A hierarchical approach to goals, which was developed in marketing and builds on some of the earlier work on values, is means-end chain theory (Reynolds and Olson, 2001; see also Huffman, Ratneshwar, and Mick, 2000). According to means-end chain theory, consumers purchase products because these products (via their need-satisfying attributes) help them attain certain values. Specifically, product attributes (both concrete attributes such as a car's horsepower and more abstract attributes such as its sportiness) have certain functional (e.g., being able to drive fast) and psycho-social (e.g., being admired by others) consequences of product use, which ultimately relate to important instrumental (e.g., being independent, imaginative) and terminal (e.g., pleasure, sense of accomplishment) values. The task of marketers is to uncover the salient means-end connections for a product category of interest (via data collection techniques such as laddering), and the resulting hierarchical value map (HVM) can then be used to segment the market and develop positioning strategies that will appeal to the chosen target segments, by emphasizing self-relevant attributes, consequences, and values and the linkages between them.

Goal structures as means-end chains. Pieters, Baumgartner and Allen (1995) extended the means-end chain approach to the modeling of consumer goal structures (see also Bagozzi and Dholakia, 1999; Baumgartner and Pieters, 2008). The notion that goal-directed behavior is structured hierarchically is a common assumption in research on goals (Austin and Vancouver, 1996). Pieters and his coauthors specifically assume that behavior is usually controlled by a basic-level goal (e.g., losing weight, purchasing a new car) at an intermediate level in the goal hierarchy (which identifies *what* the consumer is trying to do), but that in order to understand behavior more completely, one also needs to know *how* the consumer is trying to attain this goal (the subordinate goal or operation level) and *why* he or she is trying to attain this goal (the superordinate goal or motivation level). Interestingly, the goal hierarchy is closely related to two important goal features, namely, their desirability and importance on the one hand and their feasibility on the other hand. Superordinate goals are relevant for understanding why goals are desirable and important to people, whereas subordinate goals are critical for evaluating the feasibility of goal pursuit. Importantly, the three levels in the goal hierarchy are not fixed but flexible, depending on person and task

characteristics. Different consumers in different circumstances may be guided by goals at different levels in the hierarchy (ranging from wanting to avoid sugary snacks at a rather concrete level to wanting to have a more attractive body or better self-esteem at a very abstract level), and this determines which goals are subordinate and superordinate relative to the basic-level goal.

Consumers' decision making goals and purchase motives. Consumer behavior researchers have considered goals that guide decision making and the processing of other marketing stimuli (e.g., ads). For example, Bettman, Luce, and Payne (1998) proposed four goals (maximizing decision accuracy, minimizing decision effort, minimizing negative emotions during decision making, and maximizing the ease of justification of a decision) that are relevant to decision making. However, surprisingly little research has investigated the motives that underlie consumers' buying behavior (i.e., why consumers make certain buying decisions, not how they make them). Recently, Baumgartner (2010) conducted a comprehensive review of prior classifications of purchase behavior and proposed a new, empirically based typology derived from consumers' categorization of 44 purchase motives. The full typology consists of eight different purchase behaviors corresponding to the eight cells of a $2 \times 2 \times 2$ cross-classification of three orthogonal dimensions: functional versus psycho-social purchases; low versus high purchase involvement; and spontaneous versus deliberate purchases.

As might be expected, purchases that are made spontaneously and entail little effort and care (this includes casual purchases at the functional end and impulsive purchases at the psycho-social end of the continuum) are not well-defined in terms of purchase motives (i.e., buying things more or less mindlessly or based on convenience vs making unplanned or impulsive purchases). Spontaneous purchases that are higher in purchase involvement include promotional purchases (based on functional motives of getting a good deal) and exploratory purchases (based on psycho-social motives of variety, change, and curiosity).

Deliberate purchases made with little involvement comprise repetitive purchases (characterized by functional motives of loyalty, habit, routine, and familiarity) and hedonic purchases (characterized by psycho-social motives of sensory gratification and simple liking/wanting). Deliberate purchases higher in purchase involvement include extended purchase decision making (based on utilitarian motives such as logical problem solving and a concern with performance, quality, and value) and symbolic purchase behavior (involving psychological motives of wanting to express one's personality and feeling good about oneself and social motives of projecting a certain image, attaining status, and being socially accepted).

Conscious Goal Pursuit

Although they are often intertwined in practice, we distinguish two separate stages of goal pursuit: goal setting and goal striving (Bagozzi and Dholakia, 1999; Baumgartner and Pieters, 2008; Heckhausen, 1991). The discussion of goal setting will focus on the decision to pursue the basic-level goal, where the consumer commits to pursuing a goal. But, of course, goals also have to be set during goal striving at lower levels of the hierarchy, consistent with our earlier discussion about flexibility in means-end chains of goals. Once a consumer has formed an intention to pursue a particular goal (avoiding sugary snacks before dinner or eating less during the week), steps are taken to strive toward the chosen goal. According to Heckhausen (1991), goal setting is a motivational process of deciding whether or not to pursue a given goal, or choosing between conflicting goals, whereas goal striving is a volitional process focused on reaching the desired goal. We emphasize goals that are pursued deliberately in this section, but we touch upon subconscious goal pursuit too.

Goal setting. Sometimes consumers do not set goals themselves but pursue goals that are "assigned" to them (e.g., buying a wedding present using a gift registry, using someone else's shopping list to buy groceries). But often consumers set their own goals, although rarely in a social vacuum. Usually, the goal-setting process is conceptualized as some variant of subjective-expected utility or expectancy-value theory. The notion is that consumers consider

the consequences of pursuing a goal, taking into account both the desirability of the outcomes emanating from goal achievement ("part utilities" or "evaluations") and the feasibility of reaching the goal and/or the outcomes associated with the goal ("expectations" or "beliefs"). In addition, other influences are sometimes considered as well. A well-known framework, which was originally developed for explaining and predicting volitional behaviors (e.g., buying a soft drink, going out for dinner, voting for a certain candidate) but which has often been applied to goals (e.g., losing weight, donating blood), is the theory of reasoned action or TRA (Fishbein and Ajzen, 1975). According to this theory, behavior is a function of a person's intention to engage in the behavior, which in turn depends on the person's attitude toward the act and the subjective norms governing the behavior (i.e., the expectations of relevant others and the person's motivation to comply with these expectations). Attitudes toward the behavior are determined based on a person's evaluation of the consequences of engaging in the behavior weighted by the likelihood that the behavior in question will lead to these consequences. Applied to goals, the theory suggests that consumers will form a goal intention (which implies a commitment to reach the goal) if the goal is evaluated favorably by the decision maker (i.e., the outcomes associated with goal achievement are valued highly and it is deemed likely that reaching the goal will lead to these outcomes) and if relevant others are in favor of the decision maker's goal pursuit.

Bagozzi and Warshaw (1990) reformulated TRA to make it more relevant for explaining goal pursuit. This so-called theory of trying assumes that in the context of goal-directed behavior the relevant dependent variable is consumers' attempts to reach the goal in question (i.e., trying), that trying is preceded by a conscious decision to try (intention to try), and that intention to try is a function of attitude toward trying, social norms toward trying, and the frequency of past trying. Attitude toward trying depends on considerations of desirability and feasibility involving success and failure at attaining the goal as well as attitudes toward the process of goal pursuit.

Expectancy-values models have typically been applied to single behaviors and goals, but they can be extended to the choice between behaviors and goals by assuming that the behavior or goal with the highest subjective utility and/or the highest behavioral/goal intention is selected. When choosing between multiple goals, conflicts may arise. There are at least three types of conflict: approach–approach conflict (several goals are desirable, such as taking a vacation or having a baby), avoidance–avoidance conflict (one goal has to be adopted, but all choices have undesirable features, such as cleaning up the garage or painting the house) and approach–avoidance conflict (a goal has both desirable and undesirable features, such as spending one's savings when taking a vacation). Goal conflicts are omnipresent in everyday consumer behavior, but surprisingly little is known about how consumers cope with such conflicts.

Goal striving.

Goal striving with single goals. Once a consumer has formed a goal intention, which implies a commitment to attain the goal, a course of action aimed at reaching the goal has to be planned, the plan has to be implemented, progress toward goal achievement has to be monitored, and if difficulties are encountered, decisions have to be made about whether goal pursuit should be continued, reconsidered, suspended, or abandoned (see Baumgartner and Pieters, 2008).

With simple behaviors and goals, such as those usually studied with TRA, goal striving is not an issue, particularly if the necessary behaviors have been enacted frequently in the past. However, when goal achievement requires a sequence of more complex behaviors, the consumer has to develop a plan for goal pursuit. One tool for planning is mental simulation. Two types of mental simulations have been distinguished in the literature. Outcome simulation involves imagining the consequences of reaching the goal, and particularly anticipated emotional reactions to goal success and goal failure can serve a useful motivational function for energizing goal pursuit. More relevant for planning purposes are process simulations, in which the consumer imagines the steps necessary to reach the desired end.

Prior research shows that engaging in process simulations is more beneficial for goal attainment than engaging in outcome simulations (see Baumgartner and Pieters, 2008).

In a tradition strongly rooted in German psychology (Heckhausen, 1991), Gollwitzer and his associates (e.g., Gollwitzer, 1996) have argued that implementation intentions and an implemental mind-set are useful tools for enacting goal-directed behaviors. Implementation intentions are self-instructions to execute certain behaviors when a particular situation is encountered (i.e., I will do *x* when situation *y* arises). They are effective because relevant behaviors are linked to critical situations so that encountering the critical situation automatically triggers the relevant behavior or activates relevant plans in working-memory, which also increases the likelihood of enacting the intention. In a similar way, an implemental mindset has been found to facilitate the enactment of goal-relevant behaviors because instead of deliberating the pros and cons of potential courses of action (which is characteristic of a deliberative mindset), the consumer focuses single-mindedly on the execution of behaviors necessary for goal achievement, and doubts about feasibility and desirability are put aside.

Consumers need to self-regulate during goal pursuit (i.e., they have to stay focused on striving toward the goal and not be distracted by momentary temptations, and they have to be able to resume goal pursuit after an interruption). Following Baumeister (see Baumeister *et al.*, 2008, for a recent review), self-regulation involves monitoring goal progress with regard to standards and it requires self-regulatory strength. Since people's capacity for self-regulation is limited, engaging in self-regulation is ego-depleting. Various tasks have been shown to result in ego depletion, including making choices between consumer products and trading off quality and price. Furthermore, ego depletion has been demonstrated to make it more difficult for consumers to control their eating and spending, among other things, and it may lead to the use of simplistic decision strategies such as choosing options based on the attraction effect.

The final step of goal striving is the evaluation of goal achievement. If the goal is an outcome goal ("buying a Porsche sports car"), goal pursuit is terminated when the goal has been reached (i.e., the consumer has purchased the car). However, when the goal is a process goal ("learning to drive a Porsche like the pros"), goal pursuit consists of a continuous process of goal monitoring and involves attempts to either ensure goal progress or maintain adherence to appropriate standards.

If the consumer encounters difficulties during goal striving, several coping strategies are available. These include putting additional effort into goal striving, revising the initial plan, suspending the goal, or abandoning the goal. Research has shown that emotions play an important role in alerting people to those aspects of goal pursuit that require attention, and emotions are also an important consequence of goal success and goal failure (see Baumgartner and Pieters, 2008, for details).

Goal striving with multiple goals. In practice, consumers usually do not pursue single goals, but have to juggle multiple goals. One question is how they accomplish this multiple-goal pursuit over time. Will they first pursue one goal, and then move on to the next, in an orderly sequence? An integrative model of the dynamics of multiple-goal pursuit was recently proposed by Louro, Pieters, and Zeelenberg (2007). These authors argue that, when several goals compete for resources (i.e., the goals do not conflict intrinsically, but compete for the same resource pool), effort allocation to the focal goal is a joint function of the valence of goal-relevant emotions and people's proximity to the goal. When goal progress is slower (faster) than expected and people therefore experience negative (positive) emotions, less (more) effort will be allocated to the focal goal when goal attainment is distant, whereas more (less) effort will be allocated to the focal goal when goal attainment is near. Furthermore, the proposed effects are mediated by people's expectancies of success, such that either positive emotions (when the goal is distant) or negative emotions (when the goal is close) lead to moderate expectancies of success, which maximize effort allocation relative to situations of low or high expectancies of success. In other words, positive emotions have a facilitating effect on goal pursuit when people are far away from the

goal ("It is still a long way, but I'm doing well"), whereas negative emotions encourage goal pursuit when people get close to the focal goal ("It's so close, but I'm doing badly"). The reason is that these are the situations in which goal attainment is uncertain but plausible and allocation of effort promises to further goal progress. In three studies involving goals that compete for consumers' attentional and time resources, such as dieting and studying, Louro, Pieters, and Zeelenberg (2007) found support for these predictions.

Regulatory focus theory in goal pursuit. A motivational theory that has had a substantial impact on consumer research is regulatory focus theory (e.g., Pham and Higgins, 2005). The theory distinguishes two types of regulatory orientations. When people are in a promotion focus, they are concerned with accomplishments, the relevant self-guide is the ideal self (one's hopes and aspirations), and behavior is sensitive to the presence and absence of positive outcomes. In contrast, when people are in a prevention focus, they are concerned with safety, the relevant self-guide is the ought-self (one's duties and obligations), and behavior is sensitive to the presence and absence of negative outcomes. Research has shown that, consistent with these different outcome foci, promotion-oriented consumers have a preference for products signaling luxury and technical innovation, whereas prevention-oriented consumers assign greater value to safety and reliability. In addition, consumers prefer decisions for which there is a match between their regulatory orientation and the means of goal pursuit (value from fit). This occurs when a goal is pursued in an eager fashion (taking chances and avoiding missed opportunities) under a promotion orientation or in a vigilant fashion (being careful and avoiding mistakes) under a prevention orientation. Because of the dual emphasis of the theory on outcome value and value from fit, the theory is relevant for both goal setting and goal striving and has proven useful in areas as diverse as message framing and persuasion, emotional reactions to past and future decision outcomes, and cultural differences in conceptions of the self (independent vs. interdependent self-views).

Nonconscious Goal Pursuit

Thus far we have emphasized consumers' conscious pursuit of goals when a goal is set deliberately and consumers are fully aware of their attempts to strive toward the goal. Many important consumer behaviors are consistent with this process, such as when somebody decides to buy a new car or books a flight to spend the holidays with family. However, it is possible for goals to be activated outside of consumers' awareness, and even the process of goal striving may occur subconsciously. The idea of subconscious goal pursuit is an old one and in the 1950s and 1960s the term *motivation research* actually referred to the study of the hidden motives underlying purchase and consumption. Motivation research fell into disrepute and except for occasional forays into subliminal persuasion (following the journalistic report of James Vicary about the effectiveness of subliminally flashing "eat popcorn" on a movie screen), researchers showed little interest in nonconscious influences on consumer behavior. However, recently there has been a new surge of interest in automatic influences on human behavior in general and consumer behavior in particular (see Bargh, 2002; Chartrand *et al.*, 2008).

According to this perspective, processes that are functionally similar to goal setting and goal striving can occur without conscious awareness. First, cues in the environment, which have become strongly associated with certain goals through past exposure, can activate goals automatically. In laboratory studies, nonconscious activation is usually achieved through subliminal priming, in which participants are exposed to goal concepts so briefly that they are not consciously aware of the priming stimuli, or supraliminal priming, in which participants are exposed to goal concepts in the context of an ostensibly unrelated study. The real-world analog of such manipulations would be situations in which the consumer happens to get exposed to symbols that are associated with certain goals (e.g., prestige and status when seeing a store sign for Nordstrom). Prior research shows that the nonconscious activation of a goal can indeed affect consumer choices (e.g., priming prestige vs thrift increases the choice of a higher priced product).

Second, not only can goals become activated nonconsciously, but the entire process of goal pursuit can occur outside conscious awareness. The reason is that the steps necessary to reach the goal (the means of goal achievement) have been automatized and once the goal is activated, the associated action plan is enacted as well. Furthermore, this nonconscious goal pursuit exhibits many of the features characteristics of conscious goal striving. When a goal is unfulfilled, motivational strength increases over time, whereas when a goal has been achieved, motivational strength decreases. In addition, similar to conscious goals, nonconscious goals increase people's persistence when there are obstacles to goal achievement, and they encourage resumption of goal-directed behavior after interruption. For example, in the study in which either a prestige or thrift goal was primed, choice of the higher priced product as a function of the prestige prime increased when the time interval between the goal priming and the choice task was longer. Overall, there is now convincing evidence that at least relatively simple goals can be activated automatically and that the ensuing goal pursuit can also occur without consumers' conscious awareness.

Conclusion

Despite the importance of motivation and goals for understanding consumer behavior, the wanting or striving aspect of psychological functioning was not at the forefront of research for many years. However, motivation has experienced a renaissance recently, and goal concepts, both conscious and nonconscious, are now employed quite frequently in published research. We hope that this renewed emphasis will motivate additional research to deepen our understanding of consumer behavior, and that the current essay contributes to this goal.

See also *consumer desire; consumer intentions; consumer involvement; optimum stimulation level; self-regulation*

Bibliography

Austin, J.T. and Vancouver, J.B. (1996) Goal constructs in psychology: structure, process, and content. *Psychological Bulletin*, **120**, 338–375.

Bagozzi, R.P. and Dholakia, U. (1999) Goal setting and goal striving in consumer behavior. *Journal of Marketing*, **63** (Special issue), 19–32.

Bagozzi, R.P. and Warshaw, P.R. (1990) Trying to consume. *Journal of Consumer Research*, **17**, 127–140.

Bargh, J.A. (2002) Losing consciousness: automatic influences on consumer judgment, behavior, and motivation. *Journal of Consumer Research*, **29**, 280–285.

Baumeister, R.F. and Leary, M.R. (1995) The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, **117**, 497–529.

Baumeister, R.F., Sparks, E.A., Stillman, T.F. and Vohs, K.D. (2008) Free will in consumer behavior: self-control, ego depletion, and choice. *Journal of Consumer Psychology*, **18**, 4–13.

Baumgartner, H. (2010) A review of prior classifications of purchase behavior and a proposal for a new typology, in *Review of Marketing Research*, vol. 6 (ed. N.K. Malhotra), M.E. Sharpe, Armonk, NY, 3–36.

Baumgartner, H. and Pieters, R. (2008) Goal-directed consumer behavior: motivation, volition and affect, in *Handbook of Consumer Psychology* (eds C.P. Haugtvedt, P.M. Herr and F.R. Kardes), Lawrence Erlbaum, New York, pp. 367–392.

Belk, R.W., Ger, G. and Askegaard, S. (2003) The fire of desire: a multisited inquiry into consumer passion. *Journal of Consumer Research*, **30**, 326–351.

Bettman, J.R., Luce, M.F. and Payne, J.W. (1998) Constructive consumer choice processes. *Journal of Consumer Research*, **25**, 187–217.

Carver, C.S. and Scheier, M.F. (1998) *On the Self-regulation of Behavior*, Cambridge University Press, New York.

Chartrand, T.L., Huber, J., Shiv, B. and Tanner, R.J. (2008) Nonconscious goals and consumer choice. *Journal of Consumer Research*, **35**, 189–201.

Fishbein, M. and Ajzen, I. (1975) *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*, Addison-Wesley, Reading, MA.

Gollwitzer, P.M. (1996) The volitional benefits of planning, in *The Psychology of Action* (eds P.M. Gollwitzer and J.A. Bargh), Guilford, New York, pp. 287–312.

Heckhausen, H. (1991) *Motivation and Action*, 2nd edn, Springer, Berlin.

Huffman, C., Ratneshwar, S. and Mick, D.G. (2000) Consumer goal structures and goal-determination processes, in *The Why of Consumption: Perspectives on Consumer Motives, Goals, and Desires* (eds C. Huffman, R. Ratneshwar and D. Mick), Routledge, London, pp. 9–35.

Louro, M.J., Pieters, R. and Zeelenberg, M. (2007) Dynamics of multiple-goal pursuit. *Journal of Personality and Social Psychology*, **93**, 174–193.

Maslow, A.H. (1943) A theory of human motivation. *Psychological Review*, **50**, 370–396.

Pham, M. and Higgins, E.T. (2005) Promotion and prevention in consumer decision-making: State of the art and theoretical propositions, in *Inside Consumption: Frontiers of Research on Consumer Motives, Goals, and Desires*, (eds S. Ratneshwar and D.G. Mick), Routledge, New York, pp. 8–43.

Pieters, R.G.M., Baumgartner, H. and Allen, D. (1995) A means-end chain approach to consumer goal structures. *International Journal of Research in Marketing*, **12**, 227–244.

Reynolds, T.J. and Olson, J.C. (eds) (2001) *Understanding Consumer Decision Making: The Means-end Approach to Marketing and Advertising Strategy*, Lawrence Erlbaum, Mahwah, NJ.

Rokeach, M. (1973) *The Nature of Human Values*, The Free Press, New York.

Schwartz, S.H. (1992) Universals in the content and structure of values: theoretical and empirical tests in 20 countries, in *Advances in Experimental Social Psychology*, Vol. 25 (ed M. Zanna), Academic Press, New York, pp. 1–65.

consumer expertise

Eric M. Eisenstein

Introduction

Consumer learning has been a focus of research in models of consumer behavior since the early days of marketing as an academic discipline. Research on consumer knowledge and expertise is more recent (e.g., Alba and Hutchinson, 1987), but both streams of investigation share a heritage within cognitive psychology. However, within psychology, the topics of learning and expertise have had quite different research foci. Models of learning have been explored in tasks ranging from those performed by animals (e.g., pigeons and rats) to complex skills available only to humans (e.g., reading, driving, language acquisition). Many of the paradigms that have been used to investigate learning involve laboratory experiments in which subjects are "trained up" to be expert in some (usually fairly simple) task. Thus, in psychology, research into learning has largely focused on the earliest stages of learning – moving from zero or near-zero expertise to some greater level. Cognitive psychologists have traditionally viewed expertise as the end result of many years of learning and practice, and the paradigms used to identify expertise have tended toward observational studies in which learning has occurred naturally over many years rather than in the laboratory (Chi *et al.*, 1988; Shanteau, 1988a). In contrast to cognitive psychology, in consumer research neither expertise nor learning has received extensive attention, and the paradigms that have been used have not typically been the same as those used in cognitive psychology. Among other differences, in consumer research learning and expertise have generally been treated as closely related to each other.

What is expertise? Historically, there has been little agreement on the definition of expertise, which has hampered research. The layman's conception of expertise generally suggests broad-based, superior problem solving skills. This conception was shared by early researchers, who conceived experts as general problem solvers, but this formulation was short-lived, because it was soon discovered that learning experiences are bound to a specific environment, and that the superior performance of experts is tied to a domain. Initially in the concrete domain of chess, and then in many other domains, it became apparent that expertise depends critically on detailed, domain-specific knowledge, and that transfer of expertise is remarkably more difficult than expected. Therefore, when we discuss expertise, researchers are in agreement that we are really discussing human performance in a specific environment. However, there is still considerable disagreement on the definition of expertise as a construct.

Shanteau's (1988a, 1988b) Theory of Expert Competence remains the most widely accepted definition of expertise, and it provides a road map for examining consumer expertise. For the moment, let us assume the existence of a group of experts, and defer the thorny question of how one would determine who is a bona fide expert. Given the existence of at least some experts, competence will be based on five factors: domain knowledge, psychological traits, cognitive skills, decision strategies, and task characteristics. From the point of view of marketing, these characteristics can be mapped onto typical situations faced by consumers to determine the likelihood of achieving expert level consumer performance.

Adequate *domain knowledge* is a prerequisite for expert performance. It is likely that familiarity (repeated experience with the product or category) is necessary to develop adequate domain knowledge. However, it is obvious that mere memorization of facts in laundry list fashion is insufficient to achieve expert decision making performance. This is because accuracy depends critically on the linkages between stored facts, and creating and learning the linkages among stored facts is a kind of metaknowledge that is not simply declarative. Alternatively such linkages could be thought of as a separate skill requiring enrichment of the mind's ability to identify similarities and patterns, and to use that similarity function to recruit decision-relevant facts from long-term memory.

Experts require specific *cognitive skills* in order to achieve high levels of performance. The exact composition of these cognitive skills depends on the specific task. For example, the

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

task may require retrieving items from a large number of previously stored cases, the ability to perform well under pressure, or the ability to quickly recognize patterns, among other skills. Clearly, aspiring experts will have an advantage on tasks which load on these skills to the extent that their cognitive abilities match the task requirements. In addition to cognitive skills, expertise requires mastery of specific *decision strategies*, which help experts to overcome cognitive limitations. Within consumer research, considerable research effort has been devoted to the simplifying heuristics that consumers use to consider, eliminate, and choose among presented options, and how decision strategies vary with various task characteristics such as amount of information, number of attributes, number of options, and so on. Payne, Bettman and Johnson's (1993) adaptive decision maker is perhaps the classic example of this type of research (*see* CONSUMER DECISION MAKING). Research in this stream has generated robust findings that demonstrate that novice consumers frequently ignore large amounts of relevant information and are susceptible to various decision biases that arise as a result of heuristic processing. Payne *et al.*, and other researchers working within their effort-accuracy paradigm, explicitly exclude experts from consideration, and there has been little additional work done to determine what modifications to the basic theory would be necessary for expert consumers.

The role of *task characteristics* is often overlooked in the study of expertise, both in psychology and in marketing. Many have argued that characteristics of the task are critical in determining whether there will be any bona fide experts in the domain (i.e., people who are consistently superior). A common observation is that in some domains experts perform at very high levels, but in other domains, performance is not significantly different from novices. In general, task characteristics that predict good performance include repetitive tasks that are based on static, agreed-upon stimuli, with timely, veridical, feedback available, and a stationary underlying model, in which the unmodelable error in the environment is low. Task characteristics that suggest poor performance with little evidence of objective expertise include dynamic stimuli or stimuli with little agreement about which are important, nonstationary underlying processes, where feedback is either unavailable, nonveridical, delayed, or ambiguous.

Shanteau proposed that experts also possess certain *psychological traits*, such as "self-presentation – the creation and maintenance of a public image," (Shanteau, 1988b). This proposition is controversial, because there are charlatans and hucksters who are surprisingly adept at the art of self-presentation. However, surprisingly, such traits may be highly relevant when discussing consumer expertise. This is because most people have a friend or acquaintance who knows "everything" about cars, computers, stereos, food, stocks, or other common consumer goods, and one characteristic of such "consumer mavens" is that they are able to maintain (and perhaps develop) the self-presentation of expertise (*see* OPINION LEADERSHIP AND MARKET MAVENS). Consumer research has not extensively investigated the role of these consumer experts, except in thinking about them as early adopters in Bass and similar diffusion frameworks.

Applying this framework to consumer expertise, we should expect the development of expertise across a wide variety of consumer tasks. For example, many consumers buy "collectables." Some collectable purchases are priced using a typical retail take-it-or-leave-it formats. However, many collectables are sold at auction (e.g., on eBay) and others are sold at yard sales, flea markets, and other similar venues, which are not fixed price settings. Under such circumstances, we would expect that consumer expertise would develop, with some consumers able to accurately value objects within their domain of expertise. Similarly, experienced shoppers are likely to develop a reasonable sense of where to go for the cheapest product, if it is commonly purchased (though store choice is not usually determined by a single good). In other circumstances, it is unlikely that consumers will develop substantial expertise. For example, people infrequently negotiate over automobiles, little feedback is available to the consumer after the deal, and the relevant attributes change over time. Therefore, we would expect very little expertise to develop among ordinary consumers in this market.

In many domains in which expertise has been studied, expert participants are typically at the highest levels of learning or achievement within the field. By contrast, in consumer research, investigations of expertise have generally involved comparisons of more knowledgeable and less knowledgeable consumers without requiring that the more knowledgeable consumers be experts in the sense of representing the highest attainable levels of knowledge (e.g., grand masters in chess, professional judges of agricultural products, medical doctors, and meteorologists). This focus on "relative" rather than "absolute" expertise is natural because many (arguably most) important problems in consumer behavior involve the very earliest stages of learning from experience (e.g., the adoption of innovations, transitions from trial to repeat purchases, and differences between light and heavy users). Thus, the emphasis in this article is on the integration of learning and expertise on the effects of relative differences in consumer knowledge on questions of consumer behavior.

Benefits and Costs of Consumer Expertise

One reason that the definition of expertise remains open and controversial is that there are starkly conflicting research results associated with the study of experts across domains and research paradigms. Among decision scientists, the results of many carefully controlled studies of experts (as defined by credentials, schooling, years of experience, and peer consensus) paint a dismal picture of expertise, revealing a robust finding that experts are no better than novices in objective performance (see Meehl, 1954, for a seminal and comprehensive overview). These findings stand in sharp contrast to the findings of cognitive psychologists. In psychological research on experts, the robust finding is that experts dominate novices in almost every aspect of cognitive processing, such as measures of cognitive structure, memory or recall for facts, analytic processing, information use and selection, inference, problem solving strategy choice, effort and automaticity, ability to deal with complexity, and elaborative reasoning. Camerer and Johnson (1991) dubbed these contrasting results the "process-performance paradox" (i.e., experts dominate on tests of cognitive process, but fail to achieve better results). Before attempting to resolve this paradox, it will be useful to review the benefits and potential costs of expertise.

Cognitive Structure

Cognitive structure refers to the way in which factual knowledge is organized in memory. As product familiarity increases, especially from nonuser to user of a product, cognitive structure is acquired (*see* THE ROLE OF SCHEMAS IN CONSUMER BEHAVIOR RESEARCH). The most common forms of cognitive structures studied in consumer research are based on psychological research into naturally occurring categories, because product categories, consisting of brands, products, and attributes are generally learned from experience in the same way that naturally occurring taxonomic categories are learned. There are a number of widely accepted theories that are borrowed from cognitive psychology in such research. First, most product categories behave similarly to what Rosch and her colleagues (Rosch *et al.*, 1976) termed "basic level categories," meaning that objects (products) are named at this level and that people can identify objects as belonging to a basic level more quickly than those belonging to a nonbasic level. The essential characteristic of basic level categories is that, at the basic level, within-category similarity and between-category dissimilarity are maximized (e.g., the basic level category is "bird," not the subcategory "raptor" or "finch," and also not the supercategory "vertebrate" or "animal"); *see* CONSUMER CATEGORIZATION. It is easy to see how typical markets encourage classification structure at the basic level: insofar as product categories are defined as collections of substitutes that compete with each other to fulfill the same consumer needs, it is natural that such products will be quite similar to each other, and dissimilar to products that do not fulfill those needs. As expertise increases, knowledge about subordinate categories increases and information at this level is processed as efficiently as at the basic level (thus, ornithologists would be almost equally fast when classifying "raptor" or "finch"). There has been little research in

marketing to determine the extent to which the most typical products within a category are also the best products, or how this perception varies with expertise. However, marketing researchers have reached conclusions similar to those of cognitive psychologists, in that expert consumers have substantially different and more elaborate cognitive structures than novices, as well as the attendant advantages in processing speed.

Memory

One common characteristic of experts in all domains is better memory for domain-relevant facts, which is at least partially because of experts' more elaborate, better-organized, and differentiated, cognitive structures. Classic studies of chess experts showed that grand masters could remember game configurations far better than novices, but only when shown board configurations that could be obtained in real games (Chase and Simon, 1973). Similar results have been obtained in a wide variety of fields, providing strong evidence for the superiority of experts over novices in recall. Although a very small amount of this research has been specifically called consumer research (or conducted by consumer researchers), most investigations have used consumer markets or activities to define the continuum of expertise. For example, there have been extensive studies on expert versus novice players of games (e.g., chess, bridge, and go), sports (e.g., baseball, basketball, football, figure skating, and field hockey), as well as studies of hobbyists who follow these sports (i.e., people who follow baseball or basketball carefully, especially statistics related to the game). In consumer research, researchers have demonstrated the superiority of expert recall for attributes, attribute values, and the relationships between them. Studies by researchers in cognate fields have replicated the memory superiority effect in business and academic contexts, such as computer programming, mathematical proofs, physics problems, and medical diagnosis. Memory superiority, almost by definition, leads to more completely and perfectly informed consumers, and ought to improve decision making.

In general, research demonstrates that expert consumers are better able to process new information about products, are therefore more resistant to the biasing effects of advertisements, and as a result, they make objectively better decisions (however, this is not universally true, see the section Negative Consequences of Expertise).

Elaborative Thinking

In addition to better organization and memory for information acquired directly from product experiences, experts sometimes exhibit higher levels of reasoning and problem solving within their domains of expertise. The most frequent, important, and researched type of problem solving that occurs within the context of consumer expertise is the ability to infer the ultimate benefits and costs of a product based on its objective features and technical specifications, which leads to the concomitant ability to use these inferences to solve the problem of satisfying specific needs. For example, consumers are constantly faced with the task of evaluating products' attributes and determining the appropriate weights to apply to the attributes in order to maximize the consumption utility of a specific usage occasion. The research consensus is that experts are considerably more accurate than novices in making feature-to-benefit inferences, whether for themselves or when serving as an agent for another party. Expertise also reduces perceived complexity, increases the utility of the ultimate choice, and experts extract differentially more utility in complex situations than novices.

Other Positive Effects of Expertise

One surprising finding in the research on expertise is that, contrary to conventional beliefs, experts use approximately the same number of cues, or pieces of information, in their reasoning as novices. However, what information is used varies significantly between experts and novices. It appears that for many tasks, variable selection is more important than the weighting assigned to those variables. As a result, in many fields, especially consumer oriented decisions, experts perform better than novices in choosing what information to examine, where to invest time

in searching for additional information, and in making an ultimate decision.

In addition to better choices of attributes or cues to examine, experts also use substantially different problem solving strategies than novices. Among other major differences, novices tend to reason backward from a goal, whereas experts tend to classify the problem as a certain type, and then proceed forward using tools that are appropriate for the identified problem-class. This effect is particularly evident in concrete domains such as solving math, physics, or computer programming problems, as well as in games such as chess, bridge, and go. However, these games probably represent a good analogy for many types of consumer decision making. As in games, many consumer decisions are made from a strongly constrained set of options, where specific rules reduce the combinatorial space to a far smaller number (e.g., cellular phones). Hence, we should expect that expert consumers are able to reason forward rather than backward, and that this change in reasoning strategy at least partially accounts for the robust finding that expert consumers are better able to choose products appropriate to the usage situation (see the section Elaborative Thinking).

Experts have overlearned many of the skills they use. This is a double-edged sword. On the one hand, overlearning leads to automaticity, reduced effort, reduced perception of complexity, and a general reduction in effort with little or no loss of accuracy. However, automatic processing has downsides, including attentional blindness to possible environmental changes (*see* HABIT IN CONSUMER BEHAVIOR elsewhere in this volume, and the section Negative Consequences of Expertise).

Negative Consequences of Expertise

When humans believe that they have learned something, they become more confident in their answers (whether or not anything has in fact been learned in the sense of objective measures of accuracy). Unfortunately, it appears to be a nearly universal truth that the rate of increase in confidence outstrips the actual increase in accuracy (it may be noted that this does not prevent experts from being more accurate than novices, merely that there is an interaction such that they are frequently more overconfident). The correspondence between actual accuracy and confidence is called *calibration*, and most studies have found experts to be at least as poorly calibrated as novices (see Alba and Hutchinson, 2000, for a comprehensive review). This effect has been replicated in medicine, law, psychology, sports, and among undergraduates (who were assumed to have more expertise in their major than in other areas). The notable exceptions to this finding have been in examinations of professional meteorologists and world-class bridge players, leading some to speculate that large numbers of repetitions with immediate outcome feedback is necessary for calibration to be improved by experience. However, as a stylized fact, acquisition of expertise does little to reduce the overconfidence that plagues most human decision making. Given that consumer decision making environments rarely feature the type of task characteristics associated with meteorology or bridge (i.e., neither large numbers of repetitions nor immediate outcome feedback), it seems likely that consumer experts will exhibit overconfidence and lack of calibration, perhaps in greater amounts than novices.

Confidence and calibration are particularly important in the consumer domain, because confidence has been shown to be a strong predictor of consumer preferences and choice, and confident consumers are less susceptible to manipulation by advertising and other marketing techniques. The fact that confidence and learning become increasingly disassociated as expertise (and experience or familiarity) increases is therefore problematic. In particular, overconfidence among consumer experts is likely to slow the diffusion of new information into the marketplace to the extent that novice consumers rely on word of mouth and the opinions of expert consumers as a major source of prepurchase information (e.g., in high tech industries). Furthermore, the increased reach of individual experts' opinions that has resulted from Internet-based social networking is likely to make word of mouth more important in novices' decision making. Thus, social networking will likely exacerbate any negative effects associated with overconfidence.

Another possible negative consequence of expertise is that it may inhibit the spontaneous recognition of a change in the underlying environment under some circumstances (Wood and Lynch, 2002). For example, if an expert consumer has learned that attributes *a, b*, and *c*, are associated with some positive outcome, it may be more difficult for such an expert to detect the fact that *c* has become a negative, or merely less positive than it used to be. On the other hand, experts still appear to dominate novices in speed of learning, adapting to, and recognizing such changes if they are cued that a change may have occurred. Owing to the interaction, any potential disutility of expertise depends on the relative levels of cued and uncued changes in the environment – a topic that is essentially unstudied.

Explaining Expert Performance

Given the slow computational speed of the human brain, it is reasonable to ask how experts are capable of bringing relevant information, inferential techniques, and other specialized skills to bear on problems as quickly as they are observed to do.

Computation rate versus knowledge base size. Experts are called upon to process complex decision inputs and to make accurate decisions quickly. There are two general ways to approach this problem: retrieval from memory and search. Retrieval from memory is a fast and relatively effortless process for humans: conditional on a match between the current situation and a past situation that has been stored in long term memory, retrieval is rapid and automatic. If it is discovered that there is no good match in memory, however, a slower "computational" or exhaustive search through the space of possible answers ensues to discover an appropriate course of action. Search is effortful and slower, but it is the foundation of general problem solving ability. There is an obvious trade off between knowledge and search. Building an extensive database of cases and committing them to memory is time intensive, though the costs can potentially be spread out; learning procedural knowledge is frequently easier, but the actual computation is more effortful. The retrieval-search continuum represents the difference between the *cost of learning*, which is borne in the past, and the *cost of thinking*, which is incurred at the time of decision (Shugan, 1980; Eisenstein, 2002).

A robust finding is that expert performance is largely based on retrieval, not search. Experts pattern match and recall from memory, whereas novices tend to think extemporaneously. Therefore, the cost of expertise is a high cost of learning, not a cost of thinking (indeed, automaticity may reduce the cost of thinking!). This finding provides the explanation for why expertise has been found to be domain-specific and difficult to transfer to even analogous problems: the stored cases and exemplars in memory are strongly bound to a particular domain. Moreover, even for analogous problems, relevant examples must be recalled and matched, and the associative "recall key" is equally, or more strongly, bound to a domain. To provide an idea of the power of retrieval as compared to search (and therefore of experts' decision making vs novices'), we can examine the case of chess, which has been called the *drosophila* of expertise. In chess, humans can process approximately one node (essentially a potential move) per second. By contrast, computers can search nearly 200 million nodes per second. Yet, human experts compete on nearly equal terms with such computers, in spite of a hundred-million-fold difference in processing speed. In order to accomplish this feat, it has been estimated that human experts store 50 000 to 100 000 positions in memory, and simultaneously possess a highly developed sense about when each is applicable. Although the number of stored exemplars may differ by domain, it is quite likely that similar effects would be revealed in many other areas of human expertise, from appraising collectables or houses to estimating the price of cars. For consumer experts, it seems likely that similar "databases" of facts, usage occasions, outcomes, and so on, are stored by experts, whereas novice consumers must attempt to reason their way through at the time of decision. However, little research attention has been given to investigating this question.

Learning to be an expert. Expertise cannot occur without learning. As discussed previously, a large knowledge base that allows experts

to draw on a larger store of memorized facts is one difference that explains a portion of expert-novice performance differences. The large knowledge base allows experts to employ a case-based judgment strategy. Case-based strategies rely on the impressive human ability to automatically and relatively effortlessly judge similarity between cases. Using such a strategy, an expert compares the attributes of the current case to those stored in long-term memory, and the expert then reports a judgment for the current case that matches the judgment made in the stored case (or a weighted average of prior responses). This strategy has been shown to be common in fields such as law, medicine, automobile and other fault diagnostics, and chess. It is likely to be employed by experts in other domains as well (e.g., case studies are often used in MBA programs, and it is likely that business people use them as well). Case-based strategies are most useful and accurate in highly predictable environments because, in such environments, any individual case will be quite similar to the stored cases to which it is most similar. Using a case-based strategy, individual differences in the number of cases stored in long-term memory and in the veridicality of the similarity metric will underlie differences in observed expertise. The utility of a large set of stored cases begs the question of how novices accrete these cases in memory en route to becoming an expert. Cases can be observed and committed to memory either through deliberate instruction or from experience, and experiential learning may be intentional or incidental.

Instruction is one of the most common ways that people become expert in a domain. For example, someone can take a course on computer programming and master recursion, or a doctor can attend medical school and become better than a novice at diagnosis. Instruction tends to be the primary path to skilled performance in most scientific and technical fields, which makes sense because these are fields in which "practice makes perfect" and in which there is an established base of facts and relationships to be mastered. Experience is another route to skilled performance, though the literature suggests that it is a much less successful path than instruction (see Klayman, 1988 for a comprehensive review). When learning from experience, the norm is that people identify small numbers of cues and utilize simple rules for combining them. However, occasionally people do appear to learn complex rules from experience as they become expert. For example, it has been shown that there exist experts at predicting the outcome of horse races, and these experts used complex configural rules, learned from experience, that they could not articulate (and their expertise extended to making money at it).

Unfortunately for consumers, most consumer expertise is likely to be gained through experience. There are likely some exceptions where learning about a product category through a guide or magazine article is the norm. Learning from experience is more difficult than learning by instruction, and the likelihood of developing expertise from experience is greatest when the general conditions that have been identified for the development of expertise are satisfied. First, the outcome and the feedback relating to that outcome should be unambiguous. Second, the feedback about whether the prediction was correct should be immediate (delayed feedback is not particularly useful). Third, the number of experiences should be "large." For example, learning how to appraise used cars is a skill that is likely to be learned from experience. However, in order to learn to accurately price used cars, the would-be expert must observe hundreds or even thousands of used car auctions, each time comparing his predicted price to the actual selling price, updating and adding to his knowledge base. For consumers, these characteristics can be used to predict domains in which substantial expertise will develop, but these predictions are complicated by substantial heterogeneity in the availability and quality of feedback. Many consumer tasks provide ambiguous feedback, subject to a substantial delay between purchase and consumption (and therefore feedback), and a modest number of consumption opportunities. Other consumer decisions have the opposite characteristics. Moreover, for some types of feedback, such as the utility associated with using the product, it might be easy to obtain feedback that is well-suited to the development of expertise, but it might be difficult to pay appropriate attention to the feedback, as when consumption occurs in a distracting environment such as a party, or with children screaming. Because

the focus of research in marketing has been almost exclusively on studying people, rather than decision environments, we know little about the quality of feedback that is encountered by consumers.

Regardless of the quality of feedback, extensive research in both psychology and in marketing has demonstrated that raw experience or degree of familiarity is a poor predictor of expert performance. Some of this effect is owing to a subtlety in the definition of "experience." Most people think of experience as the amount of time spent "doing" an activity. For example, if someone had shopped for tennis racquets over a 20-year period, they would say that they had 20 years of experience shopping for racquets. However, experience is necessary but not sufficient to lead to improved performance. The critical link between experience and expertise appears to be the intention to learn from the experience, and at a more granular level, the precise goal associated with learning (Eisenstein and Hutchinson, 2006). This linkage occurs because learning goals define what constitutes feedback (particularly important in real-world environments in which the feedback is ambiguous), and intentionality directs attention to feedback and to discrepancies between the intended and actual outcomes. In the expertise literature, experience gained with the explicit goal of identifying discrepancies between one's mental model and the true state of the world is called *practice*, and in contrast to mere experience, time spent practicing is highly correlated with objective performance. In situations in which practice is not the norm, learning may still occur incidentally or unintentionally. This type of learning can be extremely powerful, and may lead to accurate judgments in relatively simple, deterministic domains, but it is generally not adequate for the development of expertise in most environments. Moreover, one of the hallmarks of incidental learning is the inability of subjects to articulate their reasoning process, which limits its usefulness in many decision problems.

In the development of consumer expertise, there remains an open question as far as how often consumers practice their buying skills, as opposed to simply experiencing outcomes. Certain consumer skills may be subject to substantial deliberate practice, for example searching for the lowest price on the Internet. However, some consumer activities are almost never practiced (e.g., learning attribute-price relationships in a grocery store, in which we have found that people cannot recall prices 30 seconds after purchase, much less at the time of consumption). There may also be individual differences in the degree to which identical consumer experiences are viewed as opportunities for learning, practice, and skill development (e.g., when dining out, some people are careful to taste wine with the intention of learning an attribute–quality relationship, whereas others are content simply to experience the wine).

Incentives. In laboratory studies of experts, feedback is frequently combined with an incentive to get the answer correct. In the real world, such incentives take the form of arbitrage opportunities, promotions, or other payoffs for getting the answer right. Neither the mere existence of feedback nor the addition of incentives is sufficient to guarantee the development of expertise, and in some cases, incentives reduce the rate of learning. In many consumer environments, the incentives are obvious (saved money, increased utility through superior match between product and usage). Most people, including most economists, assume that incentives will increase effort, motivation, and persistence. However, incentives are not universally helpful, and additional effort does not always result in improved performance or learning. The effects of incentives are complex and vary depending on characteristics of the task and the learning environment. In general, incentives improve performance in tasks where increased effort is likely to result in better performance, such as memory or recall tasks, clerical tasks, and motor tasks (the latter two because they are otherwise boring, and incentives help to maintain effort). Since incentives primarily affect motivation, incentives are unlikely to have an effect when intrinsic interest in the task is high. Perhaps surprisingly, incentives frequently degrade performance in tasks that require flexibility in thinking and in tasks for which it would be better to rely on automatic processes (e.g., "choking" in sports, test anxiety). Moreover,

incentives have been shown to exacerbate large classes of documented cognitive biases (e.g., representativeness, accessibility) and perceptual illusions. However, research has demonstrated that additional motivation to perform well is helpful mainly in tasks where little cognitive effort is required. Additional research has shown, however, that if incentives are too "exacting," meaning that subjects are strongly penalized for errors, then incentives reduce performance and the rate of learning. Thus, as in the case of feedback, the specific characteristics of decision making tasks are critically important in making predictions about the relationship between incentives and the development of expertise. Since incentives primarily affect motivation, one could hypothesize that incentives would be most important for the quotidian tasks consumers encounter. Perversely, however, these everyday tasks are likely to be the least important in people's lives, and also less likely to have significant incentives associated with them. Finally, in the development of consumer expertise, there are always the incentives of increased utility and decreased cost, but there has been little or no investigation of whether consumers perceive their decision milieu in these economic terms.

Conclusions and Future Research

At the outset, the question of whether there are bona fide experts was deferred, as was a satisfactory explanation for the process-performance paradox. Decision scientists have purportedly shown that experts are not superior to novices, and that experts are frequently bested in accuracy by "simple linear models." These facts are so entrenched in the decision science literature that some researchers have argued that "all anyone needs to know is which variables to look at, and then how to add" (Dawes and Corrigan, 1974).

In spite of the dismal view of experts painted by decision scientists, there are good reasons to believe that bona fide experts do exist, both in the consumer domain and in other fields that have been the frequent whipping boys of researchers (e.g., medicine). Although the evidence is mixed, taken as a whole it generally supports the superiority of experts over novices. This statement is not without caveats. The common conception of expertise is sociological – an expert is anyone agreed to be an expert. This definition is directly related to Shanteau's notion of self-presentation, and his argument that members of a field should determine who the experts are by acclamation. It may be argued that the *sine qua non* of expertise is objective accuracy. The risk with this definition is that it is tautological: experts are found only where there is demonstrable superiority in accuracy. Although this is literally true, an accuracy requirement is actually much more stringent, and more useful, because it rules out the possibility of expertise in the absence of well-defined outcomes, agreed upon by independent observers, and a gold-standard of "correctness." This requirement has the effect of eliminating the concept of expertise from many fields, because in many domains, such a standard does not exist, and it may not be possible to construct one. This is particularly true in fields for which conditions that are believed to encompass a "spectrum" of manifestations (depression, autism, Asperger's, schizophrenia), as well as for matters of taste (what does it mean for a wine or a meal to be "good"? or one operating system to be "easier to use" than another?). In both cases, what would serve as the standard of accuracy? Similarly, this definition would exclude (correctly, it may be argued) anyone from being an expert astrologer or alchemist.

In future research, the findings of lack of expert performance should be challenged. Although there have been many studies conducted by decision scientists, the vast majority have used a small set of experimental paradigms. The set of experimental paradigms should be expanded. Moreover, many (if not most) studies that have examined expert-novice differences have done so in extremely limited contexts. For example, an expert and a novice might be asked which of two options is a better match for a certain usage situation, or which of two diseases is the correct diagnosis. Such tasks minimize experts' advantages in cognitive structure, problem solving strategies, and memory because the researcher who selected the problem has already narrowed the decision space down to only a few options. By contrast, the reason that laypeople think of experts as general problem solvers is that real experts are frequently called upon to sift through an

enormous number of possibilities in order to isolate the handful that are worthy of further investigation. To use a medical analogy, once the options have been narrowed down to only flu or pneumonia, almost anyone could be taught to diagnose which is more likely. However, real doctors are instead presented with a patient who has a wet cough, fever, and feels weak. The number of pathogens that produce such symptoms is probably in the hundreds. Most will resolve without incident, but bacterial pneumonia can kill within 24 hours. Decision scientists can claim that there is no advantage to expertise, but given the symptoms above, one will seek out an experienced infectious disease doc – and when the cards are on the table, so will they.

Bibliography

Alba, J.W. and Hutchinson, J.W. (1987) Dimensions of consumer expertise. *Journal of Consumer Research*, **13** (4), 411–454.

Alba, J.W. and Hutchinson, J.W. (2000) Knowledge calibration: what consumers know and what they think they know. *Journal of Consumer Research*, **27** (2), 123–156.

Camerer, C.F. and Johnson, E.J. (1991) The process-performance paradox in expert judgment: how can experts know so much and predict so badly? in *Toward a General Theory of Expertise: Prospects and Limits*, (eds K.A. Ericsson and J. Smith), Cambridge University Press, New York.

Chase, W.G. and Simon, H.A. (1973) Perception in chess. *Cognitive Psychology*, **4** (1), 55–81.

Chi, M.T.H., Glaser, R. and Farr, M.J. (eds) (1988) *The Nature of Expertise*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Dawes, R.M. and Corrigan, B. (1974) Linear models in decision making. *Psychological Bulletin*, **81** (2), 95–106.

Eisenstein, E.M. (2002) Action-based reasoning: the cost of learning and the benefit of thinking less. University of Pennsylvania Vol. Dissertation Abstracts International. Philadelphia: Dissertation Abstracts International.

Eisenstein, E.M. and Hutchinson, J.W. (2006) Action-based learning: goals and attention in the acquisition of market knowledge. *Journal of Marketing Research*, **43** (2), 244–258.

Klayman, J. (1988) On the how and why (not) of learning from outcomes, in *Human Judgment: The SJT View* (eds B. Brehmer and C.R.B. Joyce), Elsevier Science Publihers B.V, North Holland.

Meehl, P.E. (1954) *Clinical Versus Statistical Prediction; A Theoretical Analysis and A Review of the Evidence*, University of Minnesota Press, Minneapolis, MN.

Payne, J.W., Bettman, J.R. and Johnson, E.J. (1993) *The Adaptive Decision-maker*, Cambridge University Press, Cambridge.

Rosch, E.H., Mervis, C.B., Gray, W.D., Johnson, D.M. and Boyes-Braem, P. (1976) Basic objects in natural categories, *Cognitive Psychology*, **8**, 382–439.

Shanteau, J. (1988a) Psychological characteristics and strategies of expert decision makers. *Acta Psychologica*, **68**, 203–215.

Shanteau, J. (1988b) Psychological characteristics and strategies of expert decision makers. *Acta Psychologica*, **68** (1–3), 203–215.

Shugan, S.M. (1980) The cost of thinking. *Journal of Consumer Research (pre-1986)*, **7** (2), 99.

Wood, S.L. and Lynch, J.G. (2002) Prior knowledge and complacency in new product learning. *Journal of Consumer Research*, **29** (3), 416–426.

social influence

Vladas Griskevicius and Robert B. Cialdini

Social influence involves the changing of a person's attitude or behavior through the doings of another person. Here we focus on the realm of consumer behavior and on the factors that cause one individual to comply with another's request. The starting point in ascertaining what are the most successful influence strategies was an investigation of the techniques that are most successful in professions dependent on their ability to induce compliance (e.g., salespeople, fund-raisers, advertisers, political lobbyists, cult recruiters, negotiators, and con artists). What emerged from this period of systematic observation was a list of six principles on which compliance professionals appeared to base most of their influence attempts: (i) *reciprocity* – repaying a gift, favor, or service; (ii) *consistency* – acting consistently with prior commitments; (iii) *social validation* – following the lead of similar others; (iv) *liking* – accommodating the requests of those we know and like; (v) *authority* – conforming to the directives of legitimate authorities; and (vi) *scarcity* – seizing opportunities that are scarce or dwindling in availability. A full account of the origins, workings, prevalence, and scientific evidence of these six principles is available elsewhere (Cialdini, 2009; Cialdini and Goldstein, 2004; see also Goldstein, Martin, and Cialdini, 2008). Here we provide a summary of the principles and the scientific evidence regarding how each principle functions to influence consumer behavior.

Reciprocity

Pay every debt as if God wrote the bill.
(Ralph Waldo Emerson)

One of the most powerful norms in all human cultures is that for reciprocity, which obligates individuals to return the form of behavior that they have received from another. For instance, we report liking those who report liking us; we cooperate with cooperators and compete against competitors; we self-disclose to those who have self-disclosed to us; we yield to the persuasive appeals of those who have previously yielded to one of our persuasive appeals; we try to harm those who have tried to harm us; and in negotiations, we make concessions to those who have offered concessions to us.

A *reciprocation rule* for compliance can be worded as follows: *One should be more willing to comply with a request from someone who has previously provided a favor or concession*. Under this general rule, people will feel obligated to provide gifts, favors, services, and aid to those who have given them such things first, sometimes even returning larger favors than those they have received. For example, restaurant servers who give two candies to guests along with the check increase their tips by 14.1%. A number of sales and fund-raising tactics also use this factor to advantage: the compliance professional initially gives something to the target person (e.g., a free sample), thereby causing the target to be more likely to give something in return. Often, this "something in return" is the target person's compliance with a substantial request.

The unsolicited gift, accompanied by a request for a donation, is a commonly used technique that employs the norm for reciprocity. One example is organizations sending free gifts through the mail. Including an unsolicited gift such as an individualized address labels can nearly double response rates. People often feel obligated to reciprocate even the smallest of gifts. For example, one study showed that people were more than twice as likely to fill out a lengthy survey when the request asking to complete the survey was accompanied by a handwritten Post-It note. Although such a note does not constitute a sizable gift, participants in the study recognized the extra effort and personal touch that this gesture required, and they felt obligated to reciprocate by agreeing to the request. Indeed, those who filled out the survey when it came with a handwritten sticky note returned it more promptly and gave more detailed answers.

A crucial aspect of successful reciprocity-based influence techniques involves activating the sense of obligation. The creation of obligation necessitates that the individual who desires to influence another needs to be the first to provide a gift or favor. It is noteworthy that this important aspect of reciprocity-based influence

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

techniques is often misemployed. For example, numerous commercial organizations offer donations to charity in return for the purchase of products or services – a general strategy falling under the rubric of "cause-related marketing." Yet such tit-for-tat appeals often fail to engage reciprocity properly because influence agents do not provide benefits first and then allow recipients to return the favor. The suboptimal nature of such messages can be clearly seen in the results of an experiment in hotels, in which messages that urged guests to reuse their towels were varied systematically. Messages that promised a donation to an environmental cause if guests first reused their towels were no more effective than standard pro-environmental messages. Consistent with the obligating force of reciprocity, however, a message informing guests that the hotel had already donated increased towel reuse by 26%.

Reciprocal concessions. A variation of the norm for reciprocation of favors is that for reciprocation of concessions (also called the *door-in-the-face technique*). A requester uses this procedure by beginning with an extreme request that is nearly always rejected and then retreating to a more moderate favor – the one the requester had in mind from the outset. In doing so, the requester hopes that the retreat from extreme to moderate request will spur the target person to make a reciprocal concession by moving from initial rejection of the larger favor to acceptance of the smaller one. This reciprocal concessions strategy has been successfully used in fund-raising contexts where, after refusing a larger request for donations, people become substantially more likely than before to give the average contribution. This technique has also been used effectively to solicit blood donors, whereby potential donors were first asked about participating in long-term donor program. When that request was refused, the solicitor made a smaller request for a one-time donation. This pattern of a large request (that is refused) followed by a smaller request significantly increased compliance with the smaller request, as compared to a control condition of people who were asked only to perform the smaller one-time favor (50 vs 32% compliance rate).

Related to the door-in-the-face technique is the *that's-not-all technique*, which is frequently used by sales operators. An important procedural difference between the two techniques is that, in the that's-not-all tactic, the target person does not turn down the first offer before a better second offer is provided. After making the first offer but before the target can respond, the requester betters the deal with an additional item or a price reduction. This approach has been found to be useful in selling more goods during a bake sale. One reason that this technique works appears to be the target person's desire to reciprocate for the better deal.

Social Validation

> *If you can keep your head when people all around you are losing theirs, you probably haven't grasped the situation*
> (Jean Kerr)

People frequently decide on appropriate behaviors for themselves in a given situation by searching for information as to how similar others have behaved or are behaving. This simple principle of behavior accounts for an amazingly varied army of human responses. For instance, research has shown that New Yorkers use it in deciding whether to return a lost wallet, hotel guests use it when deciding whether to reuse their towels, children with a fear of dogs use it in deciding whether to risk approaching a dog, amusement park visitors use it to decide whether to litter in a public place, audience members use it in deciding whether a joke is funny, National Park visitors use it when deciding whether to commit theft, pedestrians use it in deciding whether to stop and stare at an empty spot in the sky, and, on the alarming side, troubled individuals use it in deciding whether to commit suicide.

The *social validation rule* for compliance can be stated as follows: *One should be more willing to comply with a request for behavior if it is consistent with what similar others are thinking or doing*. Our tendency to assume that an action is more correct if others are doing it is exploited in a variety of settings. Bartenders often "salt" their tip jars with a few dollar bills at the beginning of the evening to simulate

tips left by prior customers and, thereby, to give the impression that tipping with folded money is proper barroom behavior. Church ushers sometimes prime collection baskets for the same reason and with the same positive effect on proceeds. Advertisers love to inform us when a product is the "fastest growing" or "largest selling" because they do not have to convince us directly that the product is good; they need only say that many others think so, which seems proof enough. The producers of charity telethons devote inordinate amounts of time to the incessant listing of viewers who have already pledged contributions. The message being communicated to the holdouts is clear: "Look at all the people who have decided to give; it *must* be the correct thing to do."

Social validation techniques are most likely to be effective in situations that are objectively unclear, and when there are indications of multiple similar others engaging in the particular behavior. For example, research shows that people are strongly influenced by the behavior of others when deciding whether to conserve energy in their homes. However, the influence of others' conservation behaviors increased as those others become more similar to the actual home resident: whereas other citizens of the state have an effect on conservation, behavior was more strongly influenced by the residents of the same city, and even more strongly influenced by the residents of their own neighborhood. Thus, when people are unsure, they are most likely to look to and accept the beliefs and behaviors of similar others as valid indicators of what they should believe and do themselves.

One tactic that compliance professionals use to engage the principle of social validation is called the *list technique*. This technique involves asking for a request only after the target person has been shown a list of similar others who have already complied. Studies show that when people are asked to donate money to charitable cause, those individuals who were initially shown a list of similar others who had already complied were significantly more likely to comply themselves. What's more, the longer the list, the greater was the effect.

Consistency

It is easier to resist at the beginning than at the end
(Leonardo Da Vinci)

Psychologists have long understood the strength of the consistency principle to direct human action. If we grant that the power of consistency is formidable in directing human action, an important practical question immediately arises: How is that force engaged? Psychologists think they know the answer – commitment. If a person can get you to make a commitment (that is, to take a stand, to go on record), that person will have set the stage for your consistency with that earlier commitment. Once a stand is taken, there is a natural tendency to behave in ways that are stubbornly consistent with the stand.

A *consistency rule* for compliance can be worded as follows: *after committing oneself to a position, one should be more willing to comply with requests for behaviors that are consistent with that position*. Any of a variety of strategies may be used to generate the crucial instigating commitment. One such strategy is the *foot-in-the-door technique*. A solicitor using this procedure will first ask for a small favor that is virtually certain to be granted. The initial compliance is then followed by a request for a larger, *related* favor. It has been found repeatedly that people who have agreed to the initial small favor are more willing to do the larger one, seemingly to be consistent with the implication of the initial action. For instance, home owners who had agreed to accept and wear a small lapel pin promoting a local charity were, as a consequence, more likely to contribute money to that charity when canvassed during a subsequent donation drive.

The foot-in-the-door technique is successful because performance of the initially requested action causes individuals to see themselves as possessing certain traits. For example, after taking and wearing the charity pin, subjects begin to see themselves as favorable toward charitable causes. Later, when asked to perform the larger, related favor of contributing to that charity, subjects would be more willing to do so to be

consistent with the "charitable" trait they had assigned to themselves.

Other, more unsavory techniques induce a commitment to an item and then remove the inducements that generated the commitment. Remarkably, the commitment frequently remains. For example, the *bait and switch procedure* is used by some retailers who may advertise certain merchandise (e.g., a room of furniture) at a special low price. When the customer arrives to take advantage of the special, he or she finds the merchandise to be of low quality or sold out. However, because customers have by now made an active commitment to getting new furniture at that particular store, they are more willing to agree to examine and, consequently, to buy alternative merchandise there.

A similar strategy is often employed by car dealers in the *low-ball technique*, which proceeds by obtaining a commitment to an action and *then* increasing the costs of performing the action. The automobile salesperson who "throws the low ball" induces the customer to decide to buy a particular model car by offering a low price on the car or an inflated one on the customer's trade-in. After the decision has been made (and, at times, after the commitment is enhanced by allowing the customer to arrange financing, take the car home overnight, etc.), something happens to remove the reason the customer decided to buy. Perhaps a price calculation error is found, or the used car assessor disallows the inflated trade-in figure. By this time, though, many customers have experienced an internal commitment to that specific automobile and proceed with the purchase.

Another approach to employing the consistency principle also has gained popularity among commercial compliance professionals. Rather than inducing a new commitment to their product or service, many practitioners point out existing commitments within potential customers that are consistent with the product or service being offered – a tactic called the *labeling technique*. In this way, desirable existing commitments are made more visible to the customer, and the strain for consistency is allowed to direct behavior accordingly. For example, insurance agents are frequently taught to stress to new home owners that the purchase of an expensive house reflects an enormous personal commitment to one's home and the well-being of one's family. Consequently, they argue it would only be consistent with such a commitment to home and family to purchase home and life insurance in amounts that befit the size of this commitment. Research of various kinds indicates that this sort of sensitization to commitments and to consequent inconsistencies can be effective in producing belief, attitude, and behavior change.

A more manipulative tactic than merely focusing people on their existing values is to put them in a situation where to refuse a specific request would be inconsistent with a value that people wish to be known as possessing. One such tactic is the *legitimization-of-paltry favors* (or even-a-penny-would-help) *technique*. Most people prefer to behave in ways that are consistent with a view of themselves as helpful, charitable individuals. Consequently, a fund-raiser who makes a request that legitimizes a paltry amount of aid ("could you give a contribution, even a penny would help") makes it difficult for a target to refuse to give at all; by doing so he/she risks appearing to be a very unhelpful person. Notice that this procedure does not specifically request a trivial sum; that would probably lead to a profusion of pennies and a small total take. Instead, the request simply makes a minuscule form of aid acceptable, thereby reducing the target's ability to give nothing and still remain consistent with the desirable image of a helpful individual. After all, how could a person remain committed to a helpful image after refusing to contribute when "even a penny would help"?

Liking

> *The main work of a trial attorney is to make the jury like his client*
> (Clarence Darrow)

A fact of social interaction that each of us can attest to is that people are more favorably inclined toward the needs of those they know and like. Consequently, a *liking rule* for compliance can be worded as follows: *One should be more willing to comply with the requests of liked*

individuals. Could there be any doubt that this is the case after examining the remarkable success of the Tupperware Corporation and their "home party" demonstration concept? The demonstration party for Tupperware products is hosted by an individual, usually a woman, who invites to her home an array of friends, neighbors, and relatives, all of whom know that their hostess receives a percentage of the profits from every piece sold by the Tupperware representative, who is also there. In this way, the Tupperware Corporation arranges for its customers to buy from and *for* a friend rather than from an unknown salesperson. So favorable has been the effect on proceeds ($3 million in sales per day) that the Tupperware Corporation has wholly abandoned its early retail outlets, and a Tupperware party begins somewhere every 2.7 seconds. Indeed, the success of this strategy has inspired many companies to use parties to sell their products, including cosmetics, arts-and-crafts, and even video games. Most influence agents, however, attempt to engage the liking principle in a different way: before asking a request, they get their targets to like *them*. But how do they do it? It turns out that the tactics that practitioners use to generate liking cluster around certain factors that have been shown by controlled research to increase liking.

Physical Attractiveness. Although it is generally acknowledged that good-looking people have an advantage in social interaction, research findings indicate that we may have sorely underestimated the size and reach of that advantage. There appears to be a positive reaction to good physical appearance that generalizes to such favorable trait perceptions as a talent, kindness, honesty, and intelligence. As a consequence, attractive individuals are more persuasive in terms of both changing attitudes and getting what they request. For instance, a study of Canadian Federal elections found that attractive candidates received more than two-and-a-half times the votes of unattractive ones. Equally impressive results seem to pertain to the judicial system. In a Pennsylvania study, for example, researchers rated the physical attractiveness of 74 separate male defendants at the start of their criminal trials. When, much later, the researchers checked the results of these cases via court records, they found that the better-looking men received significantly lighter sentences. In fact, the attractive defendants were twice as likely to avoid incarceration as the unattractive defendants. When viewed in the light of such powerful effects, it is not surprising that extremely attractive models are employed to promote products and services, that sales trainers frequently include appearance and grooming tips in their lessons, or that, commonly, con men are handsome and con women pretty.

Similarity. We like people who are similar to us. This fact seems to hold true whether the similarity occurs in the area of opinions, personality traits, background, or lifestyle. Not only has research demonstrated that even trivial similarities can increase liking and have profound effects on important decisions such as careers and marriage partners but also perceived attitude similarity between oneself and a stranger can automatically activate thoughts of kinship, inducing a person to behave prosocially toward that similar other. Consequently, those who wish to be liked in order to increase our compliance can accomplish that purpose by appearing similar to us in any of a wide variety of ways. For that reason, it would be wise to be careful around salespeople who just *seem* to be just like us. Many sales training programs urge trainees to "mirror and match" the customer's body posture, mood, and verbal style, as similarities along each of these dimensions have been shown to lead to positive results. Similarity in dress provides still another example. Several studies have demonstrated that we are more likely to help those who dress like us. In one study, done in the early 1970s when young people tended to dress either in "hippie" or "straight" fashion, experimenters donned hippie or straight attire and asked college students on campus for a dime to make a phone call. When the experimenter was dressed in the same way as the student, the request was granted in over two-thirds of the instances; but when the student and requester were dissimilarly dressed, a dime was provided less than half of the time. Another experiment shows how automatic our positive response to similar others can be. Marchers in a political demonstration were found not only to be more likely to sign the petition of a similarly dressed

requester but also to do so without bothering to read it first.

Compliments. Praise and other forms of positive estimation also stimulate liking. Actor Maclain Stevenson once described how his wife tricked him into marriage: "She said she liked me." Although designed for a laugh, the remark is as much instructive as it is humorous. The simple information that someone fancies us can be a bewitchingly effective device for producing return liking and willing compliance. Although there are limits to our gullibility – especially when we can be sure that the flatterer's intent is manipulative – as a rule we tend to believe praise and to like those who provide it. Evidence for the power of praise on liking comes from a study in which men received personal comments from someone who needed a favor from them. Some of the men got only positive comments, some only negative comments, and some got a mixture of good and bad. There were three interesting findings. First, the evaluator who offered only praise was liked the best. Second, this was so even though the men fully realized that the flatterer stood to gain from their liking of him. Finally, unlike the other types of comments, pure praise did not have to be accurate to work. Compliments produced just as much liking for the flatterer when they were untrue as when they were true. It is for such reasons that salespeople are educated in the art of praise. A potential customer's home, clothes, car, taste, and so on, are all frequent targets for compliments.

Cooperation. Cooperation is another factor that has been shown to enhance positive feelings and behavior. Those who cooperate toward the achievement of a common goal are more favorable and helpful to each other as a consequence. That is why compliance professionals often strive to be perceived as cooperating partners of a target person. Automobile sales managers frequently set themselves as "villains" so that the salesperson can "do battle" on the customer's behalf. The cooperative, pulling together kind of relationship that is consequently produced between the salesperson and customer, naturally leads to a desirable form of liking that promotes sales.

Scarcity

> *The way to love anything is to realize that it might be lost*
> (Gilbert Keith Chesterton)

Opportunities seem more valuable to us when they are less available. Interestingly, this is often true even when the opportunity holds little attraction for us on its own merits. Take, as evidence, the experience of Florida State University students who, like most undergraduates, rated themselves dissatisfied with the quality of their cafeteria's food. Nine days later, they had changed their minds, rating that food as significantly better than they had before. It is instructive that no actual improvement in food service had occurred between the two ratings. Instead, earlier on the day of the second rating students had learned that, because of a fire, they could not eat at the cafeteria for two weeks.

A *scarcity rule* for compliance can be worded as follows: *One should try to secure those opportunities that are scarce or dwindling*. With scarcity operating powerfully on the worth assigned to things, it should not be surprising that compliance professionals have a variety of techniques designed to convert this power to compliance. Probably the most frequently used such technique is the "limited number" tactic in which the customer is informed that membership opportunities, products, or services exist in a limited supply that cannot be guaranteed to last for long.

Related to the limited number tactic is the "deadline" technique in which an official time limit is placed on the customer's opportunity to get what is being offered. Newspaper ads abound with admonitions to the customer regarding the folly of delay: "Last three days." "Limited time offer." "One week only sale." The purest form of a decision deadline – right now – occurs in a variant of the deadline technique in which customers are told that, unless they make an immediate purchase decision, they will have to buy the item at a higher price, or they will not be able to purchase it at all. This tactic is used in numerous compliance settings. For example, a large child photography company urges parents to buy as many poses and copies

as they can afford because "stocking limitations force us to burn the unsold pictures of your children within 24 hours." A prospective health club member or automobile buyer might learn that the deal offered by the salesperson is good for that one time; should the customer leave the premises, the deal is off. One home vacuum cleaner sales company instructs its trainees to claim to prospects that "I have so many other people to see that I have the time to visit a family only once. It's company policy that even if you decide later that you want this machine, I can't come back and sell it to you." For anyone who thinks about it carefully, this is nonsense: The company and its representatives are in the business of making sales, and any customer who called for another visit would be accommodated gladly. The real purpose of the can't-come-back-again claim is to evoke the possibility of loss that is inherent in the scarcity rule for compliance.

The idea of potential loss plays a large role in human decision making. In fact, people are generally more motivated by the thought of losing something than by the thought of gaining something of equal value. For instance, home owners told about how much money they could lose from inadequate insulation are more likely to insulate their homes than those told about how much money they could save. Similar results have been obtained on college campuses where students experienced much stronger emotions when asked to imagine losses rather than gains in their romantic relationships or grade point averages.

One naturally occurring example of the consequences of increased scarcity can be seen in the outcome of a decision by county officials in Miami to ban the use and possession of phosphate detergents. Spurred by the tendency to want what they could no longer have, the majority of Miami consumers came to see phosphate cleaners as better products than before. Compared to Tampa residents, who were not affected by the Miami ordinance, the citizens of Miami rated phosphate detergents as gentler, more effective in cold water, better whiteners and fresheners, and more powerful on stains. After passage of the law, they had even come to believe that phosphate detergents poured easier than did the Tampa consumers.

Other research has suggested that in addition to commodities, limited access to information makes the information more desirable and more influential. For example, wholesale beef buyers who were told of an impending imported beef shortage purchased significantly more beef when they were informed that the shortage information came from certain "exclusive" contacts that the importer had. Apparently, the fact that the scarcity news was itself scarce made it more valued and persuasive.

Authority

Follow an expert
(Virgil)

Legitimately constituted authorities are extremely influential persons. Whether they have acquired their positions through knowledge, talent, or fortune, their positions bespeak of superior information and power. For most people, conforming to the dictates of authority figures produces genuine practical advantages. Consequently, it makes great sense to comply with the wishes of properly constituted authorities. It makes so much sense, in fact, that people often do so when it makes no sense at all.

Take, for example, the strange case of the "rectal earache" reported by two professors of pharmacy. A physician ordered eardrops to be administered to the right ear of a patient suffering pain and infection there. But instead of writing out the location "right ear" on the prescription completely, the doctor abbreviated it so that the instructions read "place in R ear." Upon receiving the prescription, the duty nurse promptly put the required number of eardrops into the patient's anus. Obviously, rectal treatment of an earache made no sense. Yet, neither the patient nor the nurse questioned it.

An *authority rule* for compliance can be worded as follows: *One should be more willing to follow the suggestions of someone who is a legitimate authority*. Authorities may be seen as falling into two categories: authorities with regard to the specific situation and more general authorities. Compliance practitioners employ techniques that seek to benefit from the power

invested in authority figures of both types. In the case of authority relevant to a specific situation, we can note how often advertisers inform their audiences of the level of expertise of product manufacturers (e.g., "Fashionable men's clothiers since 1841"; "Babies are our business, our only business"). At times, the expertise associated with a product has been more symbolic than substantive, for instance, when actors in television commercials wear physicians' white coats to recommend a product. In one famous coffee commercial, the actor involved, Robert Young, did not need a white coat, as his prior identity as TV doctor Marcus Welby, MD, provided the medical connection. It is instructive that the mere symbols of a physician's expertise and authority are enough to trip the mechanism that governs authority influence. One of the most prominent of these symbols, the bare title "Dr.," has been shown to be devastatingly effective as a compliance device among trained hospital personnel. In what may be the most frightening study we know, a group of physicians and nurses conducted an experiment that documented the dangerous degree of blind obedience that hospital nurses accorded to an individual whom they had never met, but who claimed in a phone call to be a doctor. Ninety-five percent of those nurses were willing to administer an unsafe level of a drug merely because that caller requested it.

In the case of influence that generalizes outside of relevant expertise, the impact of authority (real and symbolic) appears equally impressive. For instance, researchers have found that, when wearing a security guard's uniform, a requester could produce more compliance with requests (e.g., to pick up a paper bag in the street, to stand on the other side of a Bus Stop sign) that were irrelevant to a security guard's domain of authority. Less blatant in its connotation than a uniform, but nonetheless effective, is another kind of attire that has traditionally bespoken of authority status in our culture – the well-tailored business suit. One study found that three-and-a-half times as many people were willing to follow a jaywalker into traffic when he wore a suit and tie versus a work shirt and trousers.

Summary

An important question for anyone interested in understanding resisting or harnessing the process of influence is, "which are the most powerful principles that motivate us to comply with another's request?" We suggested that one way to assess such power would be to examine the practices of commercial compliance professionals for their pervasiveness. In other words, if compliance practitioners made widespread use of certain principles, this would be evidence for the natural power of these principles to affect everyday compliance in the realm of consumer behavior. Six principles emerged as the most popular in the repertoires of the most effective compliance pros: reciprocity, social validation, consistency, liking, scarcity, and authority. Close examination of the principles revealed broad professional usage that could be validated and explained by scientific research, indicating that these six principles engage central features of the human condition in the process of motivating compliance.

Bibliography

Cialdini, R.B. (2009) *Influence: Science and Practice*, 5th edn, Allyn & Bacon, Boston.

Cialdini, R.B. and Goldstein, N.J. (2004) Social influence: compliance and conformity. *Annual Review of Psychology*, **55**, 591–621.

Goldstein, N.J., Martin, S., and Cialdini, R.B. (2008) *Yes! 50 Scientifically Proven Ways to be Persuasive*, Free Press.

choice models

Seethu Seetharaman

A brand manager is interested in understanding market demand for his product and how marketing activities influence such demand. Since market demand is simply an aggregation of individual brand choices, numerically quantifying households' brand-choices and their dependence on marketing variables is of central importance to brand managers. Also called *discrete choice models*, brand choice models confront the problem of a household's choice of one brand from a set of mutually exclusive and collectively exhaustive brands within a product category. The model explains households' (observed) brand choices using both *observed* predictor variables (such as brands' prices, advertising activities etc.) and *unobserved* parameters (such as households' brand preferences, price sensitivities etc.), the latter being estimated using consumer choice data.

Consider the example of a household, having decided to buy a bottle of ketchup on a shopping visit, deciding on whether to buy *Heinz* or *Hunts*. The household first evaluates information about the attributes of each brand, for example quality, price, feature advertising (i.e., whether or not a brand is advertised in the newspaper that week) and end-of-aisle display (i.e., whether or not a brand is on display using a shelf talker). If one brand dominates the other along all four attributes, the household chooses the dominant brand. If it dominates the other in some attributes (e.g., price, display), but is dominated by the other brand in other attributes, (e.g., feature advertising, quality) the household's relative preferences for attributes determines which of the two brands the household will buy. A mathematical representation of how a household resolves these trade-offs between attributes in order to choose a brand is called a *brand choice model*. The household is assumed to first express the composite attractiveness of each brand (along the four attributes) using a scalar measure called *utility*. The household is then assumed to choose the brand that offers the higher utility. Since the household's utilities for brands are unknown to the marketing researcher, they must be treated as random variables. For this reason, brand choice models are also called *random utility models*. Next we describe two such random utility models – Logit and Probit.

Logit Model

We will start with a discussion of the simplest logit model, the binary logit model, and then move on to the multinomial logit model.

Binary logit – a utility theoretic exposition. Let us consider the example of the household choosing between *Heinz* and *Hunts* brands of ketchup at a shopping occasion. Suppose the household's utilities for the two brands at time t are given by U_{1t} and U_{2t} respectively. The household will choose *Heinz* at time t if $U_{1t} > U_{2t}$ and Hunts if $U_{2t} > U_{1t}$. While U_{1t} and U_{2t} are known to the household, they are unknown to the marketing researcher. The marketing researcher specifies these utilities as follows.

$$\begin{aligned} U_{1t} &= V_{1t} + \varepsilon_{1t} \\ U_{2t} &= V_{2t} + \varepsilon_{2t} \end{aligned} \tag{1}$$

where V_{1t} and V_{2t} are the deterministic components of the utilities at time t, while ε_{1t} and ε_{2t} are the random components of the utilities (that represent the effects of unobserved variables, which are known to the household but unknown to the researcher, that influence households' utilities for brands).[1] The household's probability of choosing each brand is then expressed as follows.

$$\begin{aligned} \Pr(1) &= \Pr(U_{1t} > U_{2t}) \\ &= \Pr(V_{1t} + \varepsilon_{1t} > V_{2t} + \varepsilon_{2t}) \\ &= \Pr(\varepsilon_{2t} - \varepsilon_{1t} < V_{1t} - V_{2t}) \\ \Pr(2) &= \Pr(U_{2t} > U_{1t}) \\ &= \Pr(V_{2t} + \varepsilon_{2t} > V_{1t} + \varepsilon_{1t}) \\ &= \Pr(\varepsilon_{2t} - \varepsilon_{1t} > V_{1t} - V_{2t}) \end{aligned} \tag{2}$$

The deterministic components of the utilities are specified in terms of observed brand attributes as shown below.

$$\begin{aligned} V_{1t} &= V(X_{1t}) = \alpha_1 + X_{1t}\beta \\ V_{2t} &= V(X_{2t}) = \alpha_2 + X_{2t}\beta \end{aligned} \tag{3}$$

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

where $X_{1t} = (P_{1t}D_{1t}F_{1t})$ and $X_{2t} = (P_{2t}D_{2t}F_{2t})$ are brand-specific attribute vectors at time t, where P_{it}, D_{it} and F_{it} stand for price, display and feature of brand i respectively, $\beta = (\beta_1\beta_2\beta_3)'$, is a parameter-vector associated with these attributes,[2] and $\alpha = (\alpha_1\alpha_2)'$ is a vector of brand constants. The random components ε_{1t} and ε_{2t} are assumed to be distributed *iid* Gumbel with location parameter η and scale parameter $\mu > 0$. That is,

$$F(\varepsilon_{1t}) = \mathrm{e}^{-\mathrm{e}^{-\mu(\varepsilon_{1t}-\eta)}}$$
$$F(\varepsilon_{2t}) = \mathrm{e}^{-\mathrm{e}^{-\mu(\varepsilon_{2t}-\eta)}} \tag{4}$$

This implies that $\varepsilon_{2t} - \varepsilon_{1t}$ is distributed logistic with location parameter 0 and scale parameter $\mu > 0$. That is,

$$F(\varepsilon_{2t} - \varepsilon_{1t}) = \frac{1}{1 + \mathrm{e}^{-\mu(\varepsilon_{2t}-\varepsilon_{1t})}} \tag{5}$$

Using 5 in 2 we get

$$\Pr(1) = \frac{1}{1 + \mathrm{e}^{-\mu(V_{1t}-V_{2t})}}$$
$$\Pr(2) = \frac{\mathrm{e}^{-\mu(V_{1t}-V_{2t})}}{1 + \mathrm{e}^{-\mu(V_{1t}-V_{2t})}} \tag{6}$$

which can be rewritten as

$$\Pr(1) = \frac{\mathrm{e}^{\mu V_{1t}}}{\mathrm{e}^{\mu V_{1t}} + \mathrm{e}^{\mu V_{2t}}}$$
$$\Pr(2) = \frac{\mathrm{e}^{\mu V_{2t}}}{\mathrm{e}^{\mu V_{1t}} + \mathrm{e}^{\mu V_{2t}}} \tag{7}$$

Substituting 3 in 7 yields

$$\Pr(1) = \frac{\mathrm{e}^{\mu(\alpha_1+X_{1t}\beta)}}{\mathrm{e}^{\mu(\alpha_1+X_{1t}\beta)} + \mathrm{e}^{\mu(\alpha_2+X_{2t}\beta)}}$$
$$\Pr(2) = \frac{e^{\mu(\alpha_2+X_{2t}\beta)}}{\mathrm{e}^{\mu(\alpha_1+X_{1t}\beta)} + \mathrm{e}^{\mu(\alpha_2+X_{2t}\beta)}} \tag{8}$$

Equation 8 represents the *binary logit model*. Since the parameter μ is confounded with the scale of the β's, it is usually fixed at 1. This yields the following standard version of the binary logit model.

$$\Pr(1) = \frac{\mathrm{e}^{\alpha_1+X_{1t}\beta}}{\mathrm{e}^{\alpha_1+X_{1t}\beta} + \mathrm{e}^{\alpha_2+X_{2t}\beta}}$$
$$\Pr(2) = \frac{\mathrm{e}^{\alpha_2+X_{2t}\beta}}{\mathrm{e}^{\alpha_1+X_{1t}\beta} + \mathrm{e}^{\alpha_2+X_{2t}\beta}} \tag{9}$$

The brand choice probabilities in 9 will remain unchanged if a constant is added to all brands' utilities. For this reason, the intercept term associated with one of the brands is fixed (usually at zero). This yields the following *estimable* version of the binary logit model.

$$\Pr(1) = \frac{\mathrm{e}^{\alpha_1+X_{1t}\beta}}{\mathrm{e}^{\alpha_1+X_{1t}\beta} + \mathrm{e}^{X_{2t}\beta}}$$
$$\Pr(2) = \frac{\mathrm{e}^{X_{2t}\beta}}{\mathrm{e}^{\alpha_1+X_{1t}\beta} + \mathrm{e}^{X_{2t}\beta}} \tag{10}$$

The idea of random utility goes way back to Thurstone (1927). However, a formal exposition of the logit model from random utility primitives, as laid down in this section, first appeared in McFadden (1974).

Binary logit – a regression-based exposition. The binary logit model can also be understood as a regression model, where the dependent variable y_t takes one of two values, that is, 1 or 0. The regression model looks as shown below.

$$Y_t = \pi(X_t) + \varepsilon$$

where

$$\pi(X_t) = \frac{\mathrm{e}^{\alpha_1+X_{1t}\beta}}{\mathrm{e}^{\alpha_1+X_{1t}\beta} + \mathrm{e}^{X_{2t}\beta}}$$

and

$$\varepsilon_{1t} = 1 - \pi(X_t) \text{ if } Y_t = 1$$
$$= -\pi(X_t) \text{ if } Y_t = 0 \tag{11}$$

This regression-based representation of a binary logit model is also called a *logistic regression*. There are two key differences between the conventional linear regression model and the logistic regression model: one, unlike the linear regression model where $\mathrm{E}(Y_t|X_t) = X_t\beta$ which is *linear* in X_t and *unbounded*, in the logistic regression model $\mathrm{E}(Y_t|X_t) = \pi(X_t)$, which is *S-shaped* in X_t and *bounded* between 0 and 1; two, unlike the linear regression model where ε has a *normal* distribution with mean 0 and *constant* variance σ^2, in the logistic regression model ε has a *two-point discrete* distribution with mean 0 and *heteroscedastic* variance $\pi(X_t)[1 - \pi(X_t)]$. Since the observed choice outcome Y_t has a

binomial distribution with parameter $\pi(X_t)$, that is a *logit* function of the marketing variables, this model is also called the *binomial logit* model. Next we will discuss how to estimate the parameters of the binary logit/binomial logit/logistic regression model using consumer choice data.

Maximum likelihood estimation of the binary logit model. Given n independent observations of (X_t, Y_t) the parameters of the logistic regression model are estimated by maximizing the following *likelihood* function.

$$l(\beta) = \prod_{t=1}^{n} \pi(X_t)^{Y_t}[1 - \pi(X_t)]^{1-Y_t} \quad (12)$$

which is equivalent to maximizing the following *log-likelihood* function.

$$L(\beta) = \ln[l(\beta)] = \sum_{t=1}^{n}\{Y_t \ln \pi(X_t) + (1 - Y_t)\ln[1 - \pi(X_t)]\} \quad (13)$$

Maximization involves differentiating $L(\beta)$ with respect to β and setting the resulting expressions to zero. This yields the following $(p + 1)$ *likelihood equations*.

$$\sum_{t=1}^{n}[Y_t - \pi(X_t)] = 0$$
$$\sum_{t=1}^{n} X_{tk}[Y_t - \pi(X_t)] = 0\,(k = 1, \ldots, p) \quad (14)$$

where $p = 3$ stands for the number of predictor variables in the model. Since the likelihood equations are *non-linear* in β, they require special numerical methods for their solution (unlike the linear regression model where the first-order equations are *linear* in β). Optimization modules in programming languages such as SAS, Gauss, Fortran, and so on, automatically employ these numerical methods. One such numerical method is the *Newton-Raphson* method that iterates using the formula

$$\hat{\beta}_{i+1} = \hat{\beta}_i - \left[\frac{\partial^2 L}{\partial\beta\partial\beta'}\right]^{-1}_{\hat{\beta}_i}\left[\frac{\partial L}{\partial\beta}\right]_{\hat{\beta}_i} \quad (15)$$

and converges when $|\hat{\beta}_{i+1} - \hat{\beta}_i|$ is within a predetermined tolerance level ξ. Faster methods, such as the *Davidson Fletcher Powell* (DFP) method, use approximations for the derivatives instead of directly evaluating the derivatives. Solving the likelihood equations using one of these numerical methods yields $\hat{\beta}$ *and* $\hat{\pi}(X_t)$, the maximum likelihood estimates (MLE) of β *and* $\pi(X_t)$ respectively. Before MLE became computationally easy to perform, marketing researchers employed an alternative estimation procedure to estimate the parameters of the logit model. This was based on the *logit transformation*

$$\ln\left[\frac{\pi(X_t)}{1 - \pi(X_t)}\right] = \alpha_1 + (X_{1t} - X_{2t})\beta \quad (16)$$

which suggested that one could run a linear regression of $\ln\left[\frac{\pi(X_t)}{1-\pi(X_t)}\right]$ versus $X_{1t} - X_{2t}$ in order to estimate $\hat{\beta}$. In order to do this, researchers would group observations with the same $X_{1t} - X_{2t}$ together, compute the fraction of purchase observations in each group to estimate $\hat{\pi}(X_t)$, and then run a linear regression of $\ln\left[\frac{\hat{\pi}(X_t)}{1-\hat{\pi}(X_t)}\right]$ versus $X_{1t} - X_{2t}$ across groups. For an illustration of this estimation procedure, see Jones and Zufryden (1982). In this approach, heteroscedasticity in the errors must be explicitly adjusted for. This method is now out of vogue since current statistical software packages have MLE routines built into them. Further, MLEs have desirable asymptotic properties, as discussed next.

Asymptotic properties of maximum likelihood estimators. MLEs have desirable asymptotic properties, that is, properties that can be invoked in extremely large samples (such as those available from scanner panel data) and easily derivable compared to small-sample properties. Two of these properties are described below.

1. *Consistency*. This refers to the distribution of the MLE "collapsing" on the true parameter value as the sample gets larger. It implies

the following.

$$\begin{aligned}&\lim_{n\to\infty}\Pr[\beta-\varepsilon<\hat{\beta}<\beta+\varepsilon]\\&=1\,\forall\varepsilon>0\\&\lim_{n\to\infty}E[(\hat{\beta}-\beta)^2]\\&=\lim_{n\to\infty}E[(\hat{\beta}-E\hat{\beta})^2]\\&\quad+[E\hat{\beta}-\beta]^2=0\end{aligned}\tag{17}$$

which can be rewritten as

$$\begin{aligned}&p\lim_{n\to\infty}\hat{\beta}=\beta\\&\lim_{n\to\infty}\mathrm{MSE}(\hat{\beta})=\lim_{n\to\infty}[\mathrm{Var}(\hat{\beta})\\&+\mathrm{Bias}(\hat{\beta})^2]=0\end{aligned}\tag{18}$$

where $\mathrm{MSE}(\hat{\beta})$, $\mathrm{Var}(\hat{\beta})$ and $\mathrm{Bias}(\hat{\beta})$ stand for the mean-squared-error, variance and bias associated with $\hat{\beta}$ respectively. Consistency, therefore, implies that

$$\lim_{n\to\infty}\mathrm{Var}(\hat{\beta})=0,\lim_{n\to\infty}\mathrm{Bias}(\hat{\beta})^2=0\tag{19}$$

If $\hat{\beta}$ is a consistent estimator of β, then $h(\hat{\beta})$ is a consistent estimator of $h(\beta)$ where *h(.)* is a continuous function. This is called *Slutsky's theorem*.[3] Further, if $\hat{\beta}$ is an MLE of β, then $h(\hat{\beta})$ is an MLE of $h(\beta)$ where *h(.)* is a continuous function. This greatly simplifies the algebraic derivation of the MLE of a highly non-linear function of model parameters.

2. *Best asymptotic normality (BAN)*. The distribution of the MLE can be well approximated by a normal distribution as the sample gets larger as shown below.

$$\lim_{n\to\infty}\sqrt{n}[\hat{\beta}-\beta]\sim\mathrm{N}(0,\sigma^2)\tag{20}$$

where σ^2 is called the *asymptotic variance* of the MLE. Further, the MLE is *asymptotically efficient*, that is, σ^2 tends to zero (as $n\to\infty$) faster than the variance of any other consistent estimator.[4] The asymptotic variance of $\hat{\beta}$ is equal to the *Cramer Rao Lower Bound* (i.e., the minimum variance that an unbiased estimator can have), and is estimated as shown below.

$$\mathrm{Var}(\hat{\beta})=I^{-1}(\hat{\beta})$$

where

$$\mathbf{I}(\hat{\beta})=-\begin{bmatrix}\frac{\partial^2L(\beta)}{\partial\beta_1^2}&\frac{\partial^2L(\beta)}{\partial\beta_1\partial\beta_2}&\cdots&\frac{\partial^2L(\beta)}{\partial\beta_1\partial\beta_p}\\\cdots&\frac{\partial^2L(\beta)}{\partial\beta_2^2}&\cdots&\frac{\partial^2L(\beta)}{\partial\beta_2\partial\beta_p}\\\cdots&\cdots&\cdots&\\\cdots&\cdots&\cdots&\frac{\partial^2L(\beta)}{\partial\beta_p^2}\end{bmatrix}_{\beta=\hat{\beta}}\tag{21}$$

where $\mathbf{I}(\hat{\beta})$ is called the *information matrix*. The variance of β_k is simply the kth diagonal element of $\mathrm{Var}(\hat{\beta})$. The information matrix can be written compactly as shown below.

$$\hat{\mathrm{I}}(\hat{\beta})=X'VX$$

$$V=-\begin{bmatrix}\hat{\pi}_1(1-\hat{\pi}_1)&0&\cdots&0\\\cdots&\hat{\pi}_2(1-\hat{\pi}_2)&\cdots&0\\\cdots&\cdots&\cdots&\\\cdots&\cdots&\cdots&\hat{\pi}_n(1-\hat{\pi}_n)\end{bmatrix}\tag{22}$$

where X is the $n^*(p+1)$ predictor matrix.

If $\mathrm{g}(\hat{\beta})$ is a continuous function of $\hat{\beta}$, the asymptotic variance of $\mathrm{g}(\hat{\beta})$ can be calculated as the square of the first derivative of $\mathrm{g}(\beta)$ with respect to β (evaluated at $\hat{\beta}$) times the asymptotic variance of $\hat{\beta}$. This technique, based on approximating the function $\mathrm{g}(\beta)$ as a linear function of β at $\beta=\hat{\beta}$, is called the *delta method*.

Interpreting logit parameters. In a linear regression, the coefficient β_k stands for the change in y for a unit change in X_k. In a logistic regression, β_k stands for the change in *log-odds* for a unit change in X_k, as shown below.

$$\hat{\beta}_k=\ln\left[\frac{\hat{\pi}(X_k+1)}{1-\hat{\pi}(X_k+1)}\right]-\ln\left[\frac{\hat{\pi}(X_k)}{1-\hat{\pi}(X_k)}\right]\tag{23}$$

which can also be written as

$$\hat{\beta}_k = \ln\left[\frac{\frac{\hat{\pi}(X_k+1)}{1-\hat{\pi}(X_k+1)}}{\frac{\hat{\pi}(X_k)}{1-\hat{\pi}(X_k)}}\right] = \ln[OR_k] \qquad (24)$$

where OR_k stands for the *odds-ratio* based on variable X_k. In other words, the coefficient β_k stands for the *log-odds ratio* based on the variable X_k.

Hypothesis testing. The *likelihood ratio* (LR) test is used to test hypotheses about parameters of the binary logit model, for example, testing the joint significance of a set of estimates. It is based on the following test statistic.

$$G = -2^* \ln\left[\frac{l^*}{l}\right] = -2^*[L^* - L] \sim \chi_k^2 \qquad (25)$$

where L^* is the log-likelihood of the restricted model, L is the log-likelihood of the unrestricted model, k is the number of restrictions. If one used this to test the significance of the estimated price coefficient, L^* will be the log-likelihood of the model estimated after dropping the price variable from the model, and $k = 1$. The LR test is analogous to the F-test in linear regression models.

The *Wald* test is used to test individual significance of estimated parameters. It is based on the following test statistic.

$$W = \frac{\hat{\beta}_k}{\hat{\text{SE}}(\hat{\beta}_k)} \sim t_1 \qquad (26)$$

where $\hat{\beta}_k$ is the MLE of the coefficient of the kth predictor variable and $\hat{\text{SE}}(\hat{\beta}_k)$ is the estimate of its standard error. One can also construct confidence intervals for parameters using the *Wald* statistic. The *Wald* test is analogous to the t-test in linear regression models. The difference between the *Wald* test and the LR test is that the *Wald* test employs the estimates themselves while the LR test uses the log-likelihood values instead. The *Wald* test has lower computational cost since the LR test involves the estimation of two models – restricted and unrestricted. Asymptotically, the two tests are equivalent. In fact, there is another asymptotically equivalent test called the *score test* or *Lagrange multiplier test* that uses derivatives instead.

Assessing model fit. Model fit is typically assessed on the basis of the "distance" between observed Y's and fitted $\hat{Y}$'s. Some of the popular measures of model fit are listed below.

1. *Pearson chi-square statistic.* This is given by

$$\sum_{c=1}^{C} r(Y_c, \hat{\pi}_c)^2 \sim \chi^2_{C-(p+1)} \qquad (27)$$

where

$$r(Y_c, \hat{\pi}_c) = \frac{Y_c - m_c\hat{\pi}_c}{\sqrt{m_c\hat{\pi}_c(1-\hat{\pi}_c)}} \qquad (28)$$

where Y_c stands for the number of observed positive responses with *covariate pattern c*. A *covariate pattern* refers to a single set of values for the covariates in a model, for example, price = \$1, display = 1, feature = 0. If C refers to the number of distinct values of covariates that are observed in the data ($C \leq n$), and m_c refers to the number of observations with covariate pattern c, then $\sum_{c=1}^{C} m_c = n$. If the model fits the data well, one would expect this statistic to be less than the critical value from the chi-square table.

2. *Hit rate.* Also called *classification rate*, refers to the percentage of observations that are correctly classified. It is given by

$$\frac{n - \sum_{t=1}^{n} |Y_t - \hat{Y}|}{n} \qquad (29)$$

where Y_t is the observed choice outcome at time t and $\hat{Y}_t = 1$ if $\hat{Y}_t > 0.5$. A disadvantage of the *hit rate* measure is that it depends heavily on the distribution of the true probabilities in the sample. If the true probabilities are all close to 0 or 1, most models would do equally well on this criterion. Conversely, if the true probabilities are all close to 0.5, even a good model may not do well on this criterion.

3. *Efron's (1978) R^2.* This is given by

$$1 - \frac{\sum_{t=1}^{n}(Y_t - \hat{\pi}_t)^2}{\sum_{t=1}^{n}(Y_t - \overline{Y})^2} \qquad (30)$$

This measure does not suffer from the deficiency of the *hit rate* measure, and is in the same spirit as the R^2 measure in the linear regression model. The numerator of the second term is called *sum of squared residuals*. The use of this measure cannot be defended strongly since the logit model is a heteroscedastic regression model.

4. *McFadden's (1974) Adjusted* R^2. This is given by

$$R^2_{adj} = 1 - \frac{L_p}{L_o} \tag{31}$$

where L_p is the log-likelihood of the full model and L_o is the log-likelihood of a logit model with an intercept only.

5. *Log-likelihood*. The maximized value of the log-likelihood function itself is an intuitive measure of model fit. In order to adjust for the number of parameters in the model, one can use the L-R test to test whether the change in log-likelihood in going from a restricted version of the proposed model to the proposed model is statistically significant. The L-R test requires that the two models being compared are nested models.
6. *Akaike information criterion (AIC)*. This is used to compare non-nested models; the measure is given by

$$AIC = -2^*L + 2^*k \tag{32}$$

which adjusts the log-likelihood L based on the number of parameters k.

7. *Schwarz Bayesian criterion (SBC)*. This is also used to compare non-nested models, the measure is given by

$$SBC = -2^*L + \ln n^*k \tag{33}$$

which penalizes a highly parameterized model more than the AIC if $n > e^2$.

External validation refers to the evaluation of the above-mentioned goodness-of-fit measures for a model using a *holdout sample* of observations. For example, if one used the first 80% of a household's purchase observations to fit model parameters ("calibration"), one can then evaluate hit rates, adjusted R^2 for the remaining 20% of the household's purchase observations using the estimated model parameters ("holdout validation"). Horowitz and Louviere (1993) propose a holdout validation scheme that takes into account sampling errors in predicted choices. It is based on estimating the linear regression model $Y_t = \alpha\hat{\pi}_t + \varepsilon_{jt}$, and testing the hypothesis $\hat{\alpha} = 1$.

Multinomial logit – a utility theoretic exposition. Let us consider the example of a household choosing between $\mathcal{J}$ brands of laundry detergents – *Tide*, *Wisk*, *Surf*, *Cheer*, *Bold*, *Gain*, *Era*, *Oxydol*, *Dreft*, *Ivory Snow*, and so on, at a shopping occasion. Suppose the household's utilities for the $\mathcal{J}$ brands at time t are given by $U_{1t}, \ldots, U_{\mathcal{J}t}$ respectively. The household will choose *Cheer* (say, brand i) at time t if $U_{it} > U_{jt}$ $\forall j \neq i$. The household's probability of choosing *Cheer* is then expressed as follows.

$$\Pr(i) = \Pr(U_{it} > U_{jt}, j \neq i) \tag{34}$$

which can be rewritten as

$$\Pr(i) = \Pr(U_{it} > \max_{j\neq i} U_{jt}) \tag{35}$$

which conveniently reduces the problem of making $\mathcal{J}$ pair-wise utility comparisons to making just one pair-wise comparison, that is, between U_i and $\max_{j\neq 1} U_{jt}$. If we know the distributions of the two random variables U_i and $\max_{j\neq 1} U_{jt}$, the probability in 35 can be appropriately derived.

We assume the random components in the utilities, ε_{jt}, to be distributed *iid* Gumbel with location parameter η and scale parameter $\mu > 0$. This implies that $U_{jt} = V_{jt} + \varepsilon_{jt}$ is distributed Gumbel with location parameter $\eta + V_{jt}$ and scale parameter $\mu > 0$. The distribution of $\max_{j\neq i} U_{jt}$, is derived as follows:

$$\begin{aligned} F_{\max_{j\neq i} U_{jt}}(u) &= \prod_{j\neq i} F_{U_{jt}}(u) \\ &= \prod_{j\neq i} \mathrm{e}^{-\mathrm{e}^{-\mu(u-V_{jt}-\eta)}} = \mathrm{e}^{-\sum_{j\neq i} \mathrm{e}^{-\mu(u-V_{jt}-\eta)}} \\ &= \mathrm{e}^{-\mathrm{e}^{-\mu(u-\frac{1}{\mu}\ln\sum_{j\neq i}\mathrm{e}^{\mu(V_{jt}+\eta)})}} \end{aligned} \tag{36}$$

which is distributed Gumbel with location parameter $\frac{1}{\mu}\ln\sum_{j\neq i} e^{\mu(V_{jt}+\eta)}$ and scale parameter μ. Now we can rewrite 35 as follows

$$\Pr(i) = \Pr(\max_{j\neq i} U_{jt} - U_{it} < 0) \qquad (37)$$

Since the difference between two Gumbel variates (with a common scale parameter μ) is distributed logistic whose location parameter is simply the difference between the location parameters of the two Gumbel variates, 37 can be written as follows

$$\Pr(i) = F_{\text{logistic}}(0) = \frac{1}{1 + e^{-\mu\left(\eta + V_{it} - \frac{1}{\mu}\ln\sum_{j\neq i} e^{\mu(V_{jt}+\eta)}\right)}} \qquad (38)$$

which can be written in a more convenient form as follows (setting $\eta = 0$)

$$\Pr(i) = \frac{e^{\mu V_{it}}}{\sum_{j=1}^{\mathcal{J}} e^{\mu V_{it}}} \qquad (39)$$

which represents the *multinomial logit model* (MNL). Since the parameter μ is confounded with the scale of the β's, it is usually fixed at 1. This yields the following standard version of the MNL model.

$$\Pr(i) = \frac{e^{V_{it}}}{\sum_{j=1}^{\mathcal{J}} e^{V_{it}}} = \frac{e^{\alpha_{it}+X_{it}\beta}}{\sum_{j=1}^{\mathcal{J}} e^{\alpha_{jt}+X_{jt}\beta}} \qquad (40)$$

The brand choice probabilities in 40 will remain unchanged if a constant is added to all brands' utilities. For this reason, the intercept term associated with one of the brands is fixed (usually at zero). This yields the following *estimable* version of the MNL model (where the intercept of the $\mathcal{J}$th brand has been fixed at zero).

$$\Pr(i) = \pi_i = \frac{e^{\alpha_{it}+X_{it}\beta}}{\sum_{j=1}^{\mathcal{J}-1} e^{\alpha_{jt}+X_{jt}\beta} + e^{X_{\mathcal{J}t}\beta}} \qquad (41)$$

For the first published application of the MNL model to scanner panel data on brand choices, see Guadagni and Little (1983). It is useful to note here that the MNL model is equally derivable using *deterministic utility* (instead of *random utility*) primitives. For example, one could assume that households have deterministic utilities (V_j) for brands, but have probabilistic rules that determine which brands they buy (Luce, 1959). If the probabilistic rule is the *Luce Choice Axiom*, a household's brand choice probabilities are given by

$$\Pr(i) = \pi_i = \frac{V_i}{\sum_{j=1}^{\mathcal{J}} V_j}, \qquad (42)$$

which would be identical to the MNL model in 41 if $V_i = e^{\alpha_i + X_i\beta}$.

Maximum likelihood estimation of the MNL model. Given n independent observations of (X_t, Y_t) the parameters of the MNL model are estimated by maximizing the following *likelihood* function.

$$l(\beta) = \prod_{t=1}^{n} \pi_1(X_t)^{Y_{1t}} \pi_2(X_t)^{Y_{2t}} \ldots \pi_{\mathcal{J}}(X_t)^{Y_{\mathcal{J}t}} \qquad (43)$$

which is equivalent to maximizing the following *log-likelihood* function.

$$L(\beta) = \ln[l(\beta)] = \sum_{t=1}^{n}\{Y_{1t}\ln\pi_1(X_t) + Y_{2t}\ln\pi_2(X_t) + \cdots + Y_{\mathcal{J}t}\ln\pi_{\mathcal{J}}(X_t)\} \qquad (44)$$

Maximization involves differentiating $L(\beta)$ with respect to β and setting the resulting expressions to zero. This yields the following $(J-1)^*(p+1)$ *likelihood equations*.

$$\sum_{t=1}^{n}[Y_{jt} - \pi_j(X_t)] = 0, j = 1, \ldots, \mathcal{J}-1$$

$$\sum_{t=1}^{n} X_{jkt}[Y_{jt} - \pi_j(X_t)] = 0\,(k = 1, \ldots, p), \quad j = 1, \ldots, \mathcal{J}-1 \qquad (45)$$

Since these equations are *non-linear* in β, they require special methods for their solution. One

uses the *Cramer Rao Lower Bound* to estimate the variance-covariance matrix of the estimated parameters as shown below.

$$\mathrm{Var}(\beta) = I^{-1}(\hat{\beta})$$

where

$$\mathbf{I}(\hat{\beta}) = -\begin{bmatrix} \frac{\partial^2 L(\beta)}{\partial \beta_1^2} & \frac{\partial^2 L(\beta)}{\partial \beta_1 \partial \beta_2} & \cdots & \frac{\partial^2 L(\beta)}{\partial \beta_1 \partial \beta_p} \\ \cdots & \frac{\partial^2 L(\beta)}{\partial \beta_2^2} & \cdots & \frac{\partial^2 L(\beta)}{\partial \beta_2 \partial \beta_p} \\ \cdots & \cdots & \cdots & \\ \cdots & \cdots & \cdots & \frac{\partial^2 L(\beta)}{\partial \beta_p^2} \end{bmatrix} \tag{46}$$

where $\mathbf{I}(\hat{\beta})$ is called the *information matrix*. The variance of β_k is simply the kth diagonal element of $\mathrm{Var}(\beta)$. The information matrix can be written compactly as shown below.

$$\hat{\mathbf{I}}(\hat{\beta}) = -\begin{bmatrix} \hat{\mathbf{I}}(\hat{\beta})_{11} & \hat{\mathbf{I}}(\hat{\beta})_{12} & \cdots & \hat{\mathbf{I}}(\hat{\beta})_{1\mathcal{J}-1} \\ \cdots & \hat{\mathbf{I}}(\hat{\beta})_{22} & \cdots & \hat{\mathbf{I}}(\hat{\beta})_{2\mathcal{J}-1} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \hat{\mathbf{I}}(\hat{\beta})_{\mathcal{J}-1\mathcal{J}-1} \end{bmatrix}$$

$$\hat{\mathbf{I}}(\hat{\beta})_{jj} = X' V_{jj} X$$

$$\hat{\mathbf{I}}(\hat{\beta})_{jk} = X' V_{jk} X$$

$$V_{jj} = -\begin{bmatrix} \hat{\pi}_{1j}(1-\hat{\pi}_{1j}) & 0 & \cdots & 0 \\ \cdots & \hat{\pi}_{2j}(1-\hat{\pi}_{2j}) & \cdots & 0 \\ \cdots & \cdots & \cdots & \\ \cdots & \cdots & \cdots & \hat{\pi}_{nj-1}(1-\hat{\pi}_{nj-1}) \end{bmatrix}$$

$$V_{ij} = -\begin{bmatrix} \hat{\pi}_{1j}\hat{\pi}_{1k} & 0 & \cdots & 0 \\ \cdots & \hat{\pi}_{2j}\hat{\pi}_{2k} & \cdots & 0 \\ \cdots & \cdots & \cdots & \\ \cdots & \cdots & \cdots & \hat{\pi}_{nj-1}\hat{\pi}_{nk-1} \end{bmatrix} \tag{47}$$

where X is the $n^*(p+1)$ predictor matrix.

Role of the scale parameter. The scale parameter μ of the logit model cannot be identified since it is confounded with the utility parameters β. Hence it is arbitrarily set to 1 when dealing with a given household's brand choices. Suppose we want to compare the estimated price sensitivity of two different households. The estimates are confounded with their respective scale factors. Scale factor differences (i.e., variance differences) between the two households must therefore be isolated before comparing utility parameter estimates across individuals. Otherwise, statistical tests such as the *Chow test* may lead one to falsely conclude that parameter vectors differ between households when indeed they do not. Swait and Louviere (1993) propose a way of testing equality in parameter estimates across datasets that works as follows: first, separate logit models are estimated for the two datasets, yielding maximized log-likelihood of L_1 and L_2 respectively; second, the data from the two datasets are pooled and a logit model is estimated based on the assumption of a common β, but different scale factors for the two datasets. This yields estimates of β and the ratio (μ_2/μ_1) along with the maximized log-likelihood L_μ. An LR test of L_μ versus $L_1 + L_2$ is a test of the hypothesis $\mathrm{H}_1 : \beta_1 = \beta_2 = \beta$. Therefore, even though one cannot identify the scale parameter for any given dataset, one can identify the ratio of the scale parameter for two different datasets.

Price elasticities. One useful managerial measure is the price elasticity of a brand, that is the percentage change in a brand's demand in response to a percentage change in its price. A brand's price elasticities are computed as follows

$$\mathrm{OwnElas}_i = \frac{\partial \pi_{it}}{\partial P_{it}} * \frac{P_{it}}{\pi_{it}}$$

$$\mathrm{CrossElas}_{ij} = \frac{\partial \pi_{it}}{\partial P_{jt}} * \frac{P_{jt}}{\pi_{it}} \tag{48}$$

where $\mathrm{OwnElas}_i$ and $\mathrm{CrossElas}_{ij}$ stand for the own- and cross-price elasticity (with respect to the price of brand j) of brand i, π_{it} and π_{jt} stand for the household's logit probabilities of buying brands i and j respectively at time t, P_{it} and P_{jt} stand for the prices of the ith and jth brands

respectively at time t. For a logit model it is easy to show that

$$\text{OwnElas}_i = [1 - \pi_{it}]P_{it}\beta_k$$
$$\text{CrossElas}_{ij} = -\pi_{jt}P_{jt}\beta_k \qquad (49)$$

which implies that (i) smaller brands have higher own- and cross-price elasticities than larger brands, and (ii) higher priced brands have larger own-price elasticities but smaller cross-price elasticities than lower-priced brands. These restrictions are an artifact of the logit model, and may not necessarily be appropriate for the data in hand.

The IIA property. In the MNL model, the ratio of a household's choice probabilities for any two brands is unaffected by the presence of other brands in the market. In other words,

$$\frac{P_{it}}{P_{jt}} = \frac{e^{\alpha_{it}+X_{it}\beta}}{e^{\alpha_{jt}+X_{jt}\beta}} = e^{(\alpha_{it}-\alpha_{jt})+(X_{it}-X_{jt})\beta} \qquad (50)$$

which is independent of other brands $k \neq i, j$. This property, also called *independence from irrelevant alternatives* (IIA), is unnecessarily restrictive and leads to three types of problems.[5] The first, also called the *similarity problem* (first noted by Debreu, 1960), can be understood as follows. Suppose a household chooses between two brands – *Coke* and *7-Up* – in a two-brand product market. When a new brand – *Pepsi* – is introduced to this market, the logit model would predict that the household's choice probability for the new brand will draw an equal proportion from the household's brand choice probabilities for *Coke* and *7-Up*. However, it is more reasonable to expect that *Pepsi* will draw disproportionately more from *Coke* than from *7-Up*, almost to the extent that the household's choice probability for *7-Up* will remain unchanged. The logit model cannot allow for this. The second problem, called the *dominance problem*, can be understood as follows: suppose a household has equal choice probabilities (i.e., 0.5) for *Coke* and *7-Up* when only these two brands are available in the product category, and equal choice probabilities (i.e., 0.5) for *Pepsi* and *7-Up* when only these two brands are available. Also suppose that *Pepsi* almost completely dominates *Coke* for the household when both are available, that is, when *Pepsi* and *Coke* are the only available brands, the household's choice probabilities for the two brands are 0.95 and 0.05 respectively. Such a pattern of household choices, while perfectly reasonable, also cannot be represented using a logit model. The third problem, called the *regularity problem*, can be understood as follows: suppose a household has equal choice probabilities (i.e., 0.5) for *Coke* and *7-Up* when only these two brands are available in the product category. However, the household's choice probability for *7-Up* can increase after the introduction of *Pepsi* either because the household ends up switching between cola and non-cola in a desire for variety or because the household ends up valuating *7-Up* higher after the entry of Pepsi on account of *7-Up*'s uniqueness in the product category. The logit model cannot allow for such a probability increase for *7-Up*. To summarize, therefore, the IIA restriction of the logit model leads to at least three types of problems while modeling brand choice data. The IIA restriction is an artifact of the logit model's assumption that the random components of the household's brand utilities (i.e., the ε_j's) are *iid*.

Whether or not the IIA restriction is valid for the data in hand can be assessed on the basis of the "distance" between observed and predicted shares. McFadden, Train and Tye (1977) proposed the following test statistic:

$$\frac{\sum_{m=1}^{M}(S_m - N_m\overline{P}_{jm})^2}{N\overline{P}_j} \sim \chi^2_{M-p-1} \qquad (51)$$

where the n observations in the data are first ranked from highest to lowest in terms of the predicted choice probabilities for brand j, and then sorted into M cells such that each cell contains roughly the same number of observations, S_m is the actual number of observed choices for the brand in cell m, N_m is the total number of observations in cell m, $\overline{P}_{jm}$ is the average predicted choice probability for brand j in cell m, and $\overline{P}_j$ is the average predicted choice probability for brand j in the total sample. If the critical value of the chi-square statistic is not exceeded, the IIA restriction is valid for the data in hand.

PROBIT MODEL

Multinomial probit – a utility theoretic exposition. Let us revisit the example of a household choosing between $\mathcal{J}$ brands of laundry detergents at a shopping occasion (as discussed in the section "Multinomial Logit – A Utility Theoretic Exposition"). Suppose we assume $\varepsilon_t = (\varepsilon_{1t}\varepsilon_{2t} \ldots \varepsilon_{\mathcal{J}t})'$ to be distributed $N_{\mathcal{J}}(0, \Sigma)$ we obtain the multinomial probit model (MNP). In other words, the MNP model is based on a normal distribution of errors[6] (instead of a Gumbel distribution of errors as in the MNL model), and does not assume the errors to be *iid* (as does the MNL model). This implies that $U_t = (U_{1t}U_{2t} \ldots U_{\mathcal{J}t})' = (V_{1t} + \varepsilon_{1t}V_{2t} + \varepsilon_{2t} \ldots V_{\mathcal{J}t} + \varepsilon_{\mathcal{J}t})'$ is distributed $N_{\mathcal{J}}(V_t, \Sigma)$, where $V_t = (V_{1t}V_{2t} \ldots V_{\mathcal{J}t})'$. Since the multivariate normal cumulative distribution function (cdf) does not have a closed form, the brand choice probabilities must be written as the following J-dimensional integral.

$$\Pr(i) = \pi_{it} = \int_{\varepsilon_{1t}=-\infty}^{V_{it}-V_{1t}+\varepsilon_{it}} \ldots \int_{\varepsilon_{it}=-\infty}^{\infty} \ldots \int_{\varepsilon_{\mathcal{J}t}=-\infty}^{V_{it}-V_{\mathcal{J}t}+\varepsilon_{it}} f(\varepsilon_{1t}, \ldots, \varepsilon_{it}, \ldots, \varepsilon_{\mathcal{J}t}) \, d\varepsilon_{1t} \ldots d\varepsilon_{it} \ldots d\varepsilon_{\mathcal{J}t} \tag{52}$$

which can be rewritten as the following $(J-1)$-dimensional integral

$$\Pr(i) = \pi_{it} = \int_{\delta_{1t}=-\infty}^{V_{it}-V_{1t}} \ldots \int_{\delta_{i-1t}=-\infty}^{V_{it}-V_{i-1t}} \int_{\delta_{i+1t}=-\infty}^{V_{it}-V_{i+1t}} \ldots \int_{\delta_{\mathcal{J}t}=-\infty}^{V_{it}-V_{\mathcal{J}t}} f(\delta_{1t}, \ldots, \delta_{i-1t}, \delta_{i+1t} \ldots, \delta_{\mathcal{J}t}) \, d\delta_{1t} \ldots d\delta_{i-1t} d\delta_{i+1t} \ldots d\delta_{\mathcal{J}t} \tag{53}$$

where $\delta_{jt} = \varepsilon_{jt} - \varepsilon_{it}$. By explicitly accommodating a general covariance matrix Σ of the error terms, the probit model eliminates the IIA problem associated with the logit model. Specifically, the probit model relaxes two strong assumptions of the logit model: one, that error variances are equal across brands; two, that error covariances are zero. However, relaxing *either* of these two assumptions of the logit model is sufficient to eliminate the restrictive IIA property of the logit model.[7]

Maximum likelihood estimation of the multinomial probit model. Given n independent observations of (X_t, Y_t) the parameters of the MNP model are estimated by maximizing the following *likelihood* function.

$$l(\beta) = \prod_{t=1}^{n} \pi_1(X_t)^{Y_{1t}} \pi_2(X_t)^{Y_{2t}} \ldots \pi_{\mathcal{J}}(X_t)^{Y_{\mathcal{J}t}} \tag{54}$$

which is equivalent to maximizing the following *log-likelihood* function.

$$L(\beta) = \ln[l(\beta)] = \sum_{t=1}^{n} \{ Y_{1t} \ln \pi_1(X_t) + Y_{2t} \ln \pi_2(X_t) + \ldots + Y_{\mathcal{J}t} \ln \pi_{\mathcal{J}}(X_t) \} \tag{55}$$

Since the brand choice probabilities, $\pi_i(X_t)$, do not have a closed form (see Equation 52), their evaluation requires Monte Carlo simulation (Lerman and Manski, 1981). This involves generating g random draws from the $N_{\mathcal{J}}(V_t, \Sigma)$ distribution characterizing U_t and taking the fraction of draws in which U_{it} is maximal as $\pi_i(X_t)$. As long as g is sufficiently large, the simulated value of $\pi_i(X_t)$ will be very close to its exact value (given in Equation 52). Since the maximum likelihood routine maximizes the simulated likelihood function (as opposed to the exact likelihood function), this estimation approach is also called *simulated maximum likelihood* (SML). In order to generate random draws from the $N_{\mathcal{J}}(V_t, \Sigma)$ distribution, one must first simulate a g-dimensional vector $v = (v_1 v_2 \ldots v_g)'$ of *iid* draws from the N(0,1) distribution and then use the following equation

$$U_t = Cv + V_t \tag{56}$$

where

$$\Sigma = CC' \tag{57}$$

where C is a lower-triangular matrix with positive diagonal elements. Equation 57, also called the *Cholesky decomposition*, ensures that Σ is positive definite (which is required by the statistical regularity conditions imposed on the MNP model).

The brand choice probabilities, $\pi_i(X_t)$, can also be evaluated using a numerical approximation approach proposed by Clark (1961). In this approach, the brand choice probability is written as

$$\Pr(i) = \Pr[\max_{j \neq i} (U_{jt} - U_{it}) < 0] \quad (58)$$

where one approximates the univariate cdf of the random variable $\max_{j \neq i}(U_{jt} - U_{it})$ using recursive formulas for its mean and variance, in order to obtain the desired probability. While this approach is much quicker than Monte Carlo simulation, its bias depends on the (unknown) covariance structure of the brands' random utilities and cannot be controlled by increasing the number of observations.

Smooth recursive conditioning (SRC) simulator. The Monte Carlo simulation procedure of Lerman and Manski (1981) employs a frequency-based simulator of brand choice probabilities. Such a simulator may be discontinuous in model parameters, which may impede numerical optimization, unless the number of draws is kept very high. Borsch-Supan and Hajivassiliou (1993) employ a Monte Carlo simulation procedure to evaluate $\pi_i(X_t)$ that is computationally superior to the frequency-based simulation procedure of Lerman and Manski (1981). Also called the *smooth recursive conditioning* (SRC) simulator,[8] their Monte Carlo procedure is based on evaluating the (J − 1)-variate integral given in Equation 53. First Equation 52 is rewritten as shown below.

$$\begin{bmatrix} -\infty \\ \cdot \\ \cdot \\ -\infty \\ -\infty \\ \cdot \\ \cdot \\ -\infty \end{bmatrix} \leq \begin{bmatrix} \delta_{1t} \\ \cdot \\ \cdot \\ \delta_{i-1t} \\ \delta_{i+1t} \\ \cdot \\ \cdot \\ \delta_{\mathcal{J}t} \end{bmatrix} \leq \begin{bmatrix} V_{it} - V_{1t} \\ \cdot \\ \cdot \\ V_{it} - V_{i-1t} \\ V_{it} - V_{i+1t} \\ \cdot \\ \cdot \\ V_{it} - V_{\mathcal{J}t} \end{bmatrix} \quad (59)$$

which can be written in vector-form as shown below.

$$a \leq A\varepsilon_t \leq b_t,$$

$$\mathbf{A}_{(\mathcal{J}-1)*\mathcal{J}} = \begin{bmatrix} 1 & & & -1 & & & \\ & 1 & & -1 & & & \\ & & \dots & -1 & & & \\ & & & \dots & & & \\ & & & -1 & \dots & & \\ & & & -1 & & 1 & \\ & & & -1 & & & 1 \end{bmatrix}$$

$$\varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \cdot \\ \cdot \\ \varepsilon_{it} \\ \cdot \\ \cdot \\ \varepsilon_{\mathcal{J}t} \end{bmatrix} \quad (60)$$

where

$$a = \begin{bmatrix} -\infty \\ \cdot \\ \cdot \\ -\infty \\ -\infty \\ \cdot \\ \cdot \\ -\infty \end{bmatrix}, b = \begin{bmatrix} V_{it} - V_{1t} \\ \cdot \\ \cdot \\ V_{it} - V_{i-1t} \\ V_{it} - V_{i+1t} \\ \cdot \\ \cdot \\ V_{it} - V_{\mathcal{J}t} \end{bmatrix} \quad (61)$$

Therefore, the necessary random draws are from a *truncated multivariate normal* distribution as shown below.

$$A\varepsilon_t \sim N_{\mathcal{J}-1}(0, A\Sigma A')$$

such that

$$a \leq A\varepsilon_t \leq b_t \quad (62)$$

which is equivalent to the following truncated normal distribution (Geweke, 1991).

$$\mathrm{e}_t \sim N(0, I)$$

such that

$$a \leq L\mathrm{e}_t \leq b_t \quad (63)$$

where L is the lower-triangular matrix of the Cholesky decomposition

$$LL' = A\Sigma A' \tag{64}$$

This renders the Monte Carlo simulation to be recursive as shown below[9].

$$\begin{aligned} &\mathrm{e}_1 \sim N(0,1)\ st\ a_1 \le l_{11}\mathrm{e}_1 \le b_1 \\ &\mathrm{e}_2 \sim N(0,1)\ st\ a_2 \le l_{21}\mathrm{e}_1 + l_{22}\mathrm{e}_2 \le b_2 \\ &\mathrm{e}_3 \sim N(0,1)\ st\ a_3 \le l_{31}\mathrm{e}_1 + l_{32}\mathrm{e}_2 + l_{33}\mathrm{e}_3 \le b_3 \end{aligned} \tag{65}$$

and so on.

This means that sequential sampling from univariate truncated normal distributions is sufficient to simulate the desired brand choice probability. Sampling from a univariate truncated normal distribution is achieved using the integral transform theorem as shown below.

$$\begin{aligned} &Z = \Phi^{-1}[\{\Phi(b) - \Phi(a)\}U + \Phi(a)] \\ &U \sim \mathrm{Unif}(0,1) \end{aligned} \tag{66}$$

Identifiability of probit model parameters. The identifiability of the parameters of an MNP model cannot be verified directly using the brand choice probabilities (as in the MNL model) since these probabilities do not have closed form expressions. An obvious check for identifiability is to ensure that there is no change of variable that can reduce the number of parameters in V_t and Σ. Identifiability requires one of the brand constants to be restricted to 0 (as in the MNL model), that is, to fix the location. Identifiability also requires that one brand's error variance is restricted to 1, while its error covariances are restricted to 0 (for the same reason that μ is restricted to be 1 in the MNL model), that is, to fix the scale. In more complicated versions of the MNP, one way to check for model identifiability is to check whether the Hessian matrix is negative-definite at convergence.[10] However, this is neither a necessary nor sufficient condition for model identifiability. In fact, even if the MNP model is fully identified, the log-likelihood function is not necessarily globally concave. The only way to check for global maxima is to first identify whether multiple local maxima exist (using different starting values for the parameters) and then pick the largest among the identified local maxima as the global maximum.

For marketing applications of the MNP model to scanner panel data on brand choices, see Papatla and Krishnamurthi (1992) and Chintagunta (1992). An alternative way to relax the IIA restriction is to allow the model coefficients (price, display, feature, and brand constants) in an MNL model to be random. If they follow a joint multivariate distribution, the IIA restriction is removed (even if the brand-specific errors are *iid*). Such an approach has been adopted in an independent probit framework by Hausman and Wise (1978).

Other Non-IIA Models of Brand Choice[11]

Because of the computational burden and identification problems associated with the MNP model, marketing researchers have proposed alternative brand choice models that have tractable, closed form expressions, yet avoid the IIA restriction of the MNL model. We discuss some of these models below. All of these models can be estimated using maximum likelihood.

Generalized extreme value (GEV) model. A generalization of the MNL model, first proposed by McFadden (1978), can be written as shown below.

$$\pi_t(i) = \frac{\mathrm{e}^{V_{it} + \ln G_i(\mathrm{e}^{V_{1t}}, \ldots, \mathrm{e}^{V_{\mathcal{J}t}})}}{\sum_{j=1}^{\mathcal{J}} \mathrm{e}^{V_{jt} + \ln G_j(\mathrm{e}^{V_{1t}}, \ldots, \mathrm{e}^{V_{\mathcal{J}t}})}} \tag{67}$$

where $G_i = \frac{\partial G}{\partial y_i}$, and G(.) is a function with the following properties.

1. G is non-negative,
2. G is homogeneous of degree $\mu > 0$, that is, $\mathrm{G}(\alpha \mathrm{y}_1, \ldots, \alpha \mathrm{y}_{\mathcal{J}}) = \alpha^{\mu}\, \mathrm{G}(\mathrm{y}_1, \ldots, \mathrm{y}_{\mathcal{J}})$,
3. $\lim_{y_i \to \infty} \mathrm{G}(\mathrm{y}_1, \ldots, \mathrm{y}_{\mathcal{J}}) = \infty$, for $i = 1, \ldots, \mathcal{J}$,
4. $\frac{\partial^l G}{\partial y^l} \ge 0$ if l is odd, and $\frac{\partial^l G}{\partial y^l} \le 0$ if l is even.

The GEV model is consistent with random utility maximization and can be derived by

assuming the distribution of the random components of the utilities to be $F(.) = e^{-G(e^{-\varepsilon_1},\ldots,e^{-\varepsilon_{\mathcal{J}}})}$. If $G(e^{V_1},\ldots,e^{V_{\mathcal{J}}}) = \sum_{j=1}^{\mathcal{J}} e^{V_j}$, the GEV reduces to the MNL model. If $G(e^{V_1},\ldots,e^{V_{\mathcal{J}}}) = \sum_{a=1}^{A}\left(\sum_{j\in Da} e^{V_j}\right)^{\mu}$, where the choice set $(1, \ldots, \mathcal{J})$ is partitioned into A non-overlapping subsets, the GEV reduces to the *nested logit* model (described in the next section). In fact, Equation 67 shows that the GEV model extends the MNL model by making the utility of each brand depend not only on its own attributes, but also on other brands' attributes.

Nested logit model. This model allows brands to share unobserved (and observed) attributes, which induces correlations in a household's random utilities across brands. Suppose brand i can be defined as a bundle (a, b) of two attributes.

$$U_{it} = U_{abt} = V_{at} + V_{bt} + V_{abt} + \varepsilon_{at} + \varepsilon_{bt} + \varepsilon_{abt} \tag{68}$$

where V_{at} is the deterministic component of utility common to all brands that share attribute a, V_{bt} is the deterministic component of utility common to all brands that share attribute b, ε_{at} is the random component of utility due to attribute a (i.e., varies only across brands with different a), ε_{bt} is the random component of utility due to attribute b (i.e., varies only across brands with different b), and ε_{abt} is the random component of utility that varies across all brands (i.e., the "usual" random component). Assuming that the three random components – ε_{at}, ε_{bt} and ε_{abt} – are independent of each other, have zero means and variances given by σ_a^2, σ_b^2, and σ^2 respectively, one can show that

$$\begin{aligned} \mathrm{cov}(U_{ab}, U_{ab'}) &= \sigma_a^2 \\ \mathrm{cov}(U_{ab}, U_{a'b}) &= \sigma_b^2 \end{aligned} \tag{69}$$

that is, utilities across brands are correlated.[12]

The estimable version of the nested logit model is obtained by setting one of the two attribute-specific random components to zero, for example, $\varepsilon_{bt} = 0$. Further, one assumes that $\varepsilon_{\mathrm{abt}}$ is distributed *iid* Gumbel $(0, \mu)$, and ε_{at} is distributed so that Max $_{b\in Da} U_{\mathrm{abt}}$ is distributed *iid* Gumbel $(0, \mu')$, where Da refers to the set of all brands that share attribute a. This results in the following marginal probability of a household buying a brand with attribute a.

$$\pi_{at} = \frac{e^{\left(V_{at} + \frac{1}{\mu}\ln \sum_{b\in Da} e^{(V_{bt}+V_{abt})\mu}\right)\mu'}}{\sum_{a'} e^{\left(V_{a't} + \frac{1}{\mu}\ln \sum_{b\in Da} e^{(V_{bt}+V_{a'bt})\mu}\right)\mu'}} \tag{70}$$

where $\frac{\mu'}{\mu} = \sqrt{1 - \mathrm{corr}(U_{ab}, U_{ab'})} \leq 1$. The household's *conditional* probability of buying a brand with attribute b, given choice of a brand with attribute a, is given by

$$\pi_{(b|a)t} = \frac{e^{(V_{bt}+V_{abt})\mu}}{\sum_{b'\in Da} e^{(V_{b't}+V_{ab't})\mu}} \tag{71}$$

Since only the ratio of the two scale parameters is identified, μ is usually restricted to be 1. The estimable version of the nested logit model, therefore, becomes

$$\begin{aligned} \pi_{it} = \pi_{abt} &= \pi^*_{at}\pi_{(b|a)t} \\ &= \frac{e^{\left(V_a + \ln \sum_{b\in Da} e^{V_b+V_{ab}}\right)\mu'}}{\sum_{a'} e^{\left(V_{a'} + \ln \sum_{b\in Da} e^{V_b+V_{a'b}}\right)\mu'}} \\ &* \frac{e^{V_{bt}+V_{abt}}}{\sum_{b'\in Da} e^{V_{b't}+V_{ab't}}} \end{aligned} \tag{72}$$

where $\mu' \leq 1$. If $\mu' = 1$ this model reduces to an MNL model. The nested logit model can be understood using a *tree-structure* where the upper level represents the household's choice of attribute a, and the lower level represents the household's choice of attribute b conditional on the choice of attribute a. With three attributes, the estimable nested logit model looks as shown below.

$$\pi_{it} = \pi_{abct} = \pi^*_{at}\pi^*_{(b|a)t}\pi_{(c|ab)t}$$

$$= \frac{e^{(V_{at}+V'_{at})\mu''}}{\sum_{a'\in D} e^{(V_{a't}+V'_{a't})\mu''}} * \frac{e^{(V_{bt}+V_{abt}+V'_{abt})\mu'}}{\sum_{b'\in Da} e^{(V_{b't}+V_{ab't}+V'_{ab't})\mu'}} * \frac{e^{V_{ct}+V_{bct}+V_{act}+V_{abct}}}{\sum_{c'\in Dab} e^{V_{c't}+V_{bc't}+V_{ac't}+V_{abc't}}} \quad (73)$$

where

$$V'_{abt} = \ln\left(\sum_{c'\in Dab} e^{V_{c't}+V_{bc't}+V_{ac't}+V_{abc't}}\right)$$

$$V'_{at} = \frac{1}{\mu'}\ln\left(\sum_{b'\in Da} e^{(V_{b't}+V_{ab't}+V'_{ab't})\mu'}\right)$$

$$\mu' \le 1,\ \mu'' \le \mu' \quad (74)$$

If $\mu'' = \mu'$ this model reduces to the MNL model. While the nested logit model presents a parsimonious way of modeling inter-brand correlations, determining the sequence of attributes in the tree is a subjective decision. Some authors prescribe the estimation of nested logit models with all possible orders of attributes, and pick whichever best describes the observed data. For an application of the nested logit model to scanner panel data on brand choices, see Kannan and Wright (1991).

Batsell and Polking (1985) model. This model allows the ratio of a household's choice probabilities for brands i and j to depend on other brands in the household's choice set. Specifically,

$$\beta_{ij}^{A} = \ln\left(\frac{\pi_i^A}{\pi_j^A}\right) = \sum_{I\subset[A-\{i,j\}]} \alpha_{ij}^{I} \quad (75)$$

where A is the household's choice set, that is, a subset of the full set of brands {1,...,J}. This implies that the ratio of the brands' choice probabilities is a function of the brands in the choice set A. For example, if $\mathrm{A} = \{\mathrm{i,j}\}$, $\mathrm{I} = (\{\emptyset\})$ and $\beta_{ij}^{ij} = \alpha_{ij}^{\emptyset}$. If $\mathrm{A} = \{\mathrm{i,j,k}\}$, $\mathrm{I} = (\{\emptyset\},\{\mathrm{k}\})$ and $\beta_{ij}^{\mathrm{ijk}} = \alpha_{ij}^{\emptyset} + \alpha_{ij}^{k}$. If $\mathrm{A} = \{\mathrm{i,j,k,l}\}$, $\mathrm{I} = (\{\emptyset\},\{\mathrm{k}\},\{\mathrm{l}\},\{\mathrm{k,l}\})$ and $\beta_{ij}^{\mathrm{ijkl}} = \alpha_{ij}^{\emptyset} + \alpha_{ij}^{k} + \alpha_{ij}^{l} + \alpha_{ij}^{kl}$. The estimable version of this model is given below.

$$P_A(i) = \frac{\prod_{I\subset A} e^{s_i^I}}{\sum_{k\in A}\prod_{I\subset A} e^{s_k^I}} \quad (76)$$

where

$$s_i^I = 0 \text{ for } i \notin I$$

$$\sum_{i\notin I} s_i^I = u_I \quad (77)$$

where u_I is a pre-assigned number. If $s_i^I = 0$ for $I \neq \emptyset$, this model reduces to the MNL model. Once the estimates for $s_i^{I'}s$ are obtained, the $\alpha_{ij}^{I'}s$ are given by

$$\alpha_{ij}^{I} = s_i^{iI} + s_i^{ijI} - s_j^{jI} - s_j^{ijI} \text{ if } I\cap i,j = \emptyset \quad (78)$$

While this model relaxes the IIA restriction of the MNL model in a more flexible manner than the nested logit model, it involves the estimation of a large number of parameters. Further, unlike the GEV or nested logit models, this model cannot be derived from random utility maximization.

Alternative approaches. The *elimination-by-aspects* (EBA) model of (Tversky, 1972) works on the idea that a household sequentially screens brands based on attributes in order to arrive at an optimal brand choice. This model does not have the IIA property. A *consideration set model* explains a household's brand choice as an outcome of a two-stage process, where the first stage involves the formation of a consideration set, and the second stage involves the optimal choice of brands within the consideration set (see, for example, Roberts and Lattin, 1991). Such a two-stage model also does not have the IIA property. In general, a brand choice model that is consistent with two-stage decision-making on the part of the household eliminates the IIA restriction. Similarly, any brand choice model that allows a brand's utility to depend on the attributes of competing brands (for example, the GEV model) ends up eliminating the IIA restriction. Models that obey this property are called *mother*

logit models. While a large number of non-IIA brand choice models have been proposed in the marketing literature, the most popular (in terms of the number of academic papers published on the model) are the multinomial probit and the nested logit models. Among these two, given the sharp rise in computational power in recent years and the natural appeal of the normal distribution to statisticians and econometricians, the multinomial probit model has come to be accepted as the *de facto standard* for brand choice models. Another reason for the increasing popularity of the multinomial probit model is that it provides a more flexible framework to accommodate the effects of unobserved heterogeneity, state dependence, and so on.

Endnotes

[1] It is on account of these random components that brand choice models are also called *random utility models*.

[2] It is useful to note that this parameter vector is common across brands.

[3] This is what renders asymptotic properties of estimators to be much more convenient to work with than small-sample properties.

[4] If $\hat{\beta}$ is a vector, the asymptotic variance of the MLE is a matrix $\sum$, and asymptotic efficiency implies that $(\sum^* - \sum)$ is *non-negative-definite*, where $\sum^*$ is the asymptotic variance of any other consistent estimator of β.

[5] The exposition of these problems borrows heavily from Batsell and Polking (1985).

[6] If one viewed the error term to be the sum of a large number of unobserved but independent components, by the *central limit theorem* the error term will tend to be normal.

[7] The probit model solves the *similarity problem* associated with the IIA property (see section on the IIA property), however, the *dominance* and *regularity* problems still remain.

[8] Also called the Geweke-Hajivassiliou-Keane (GHK) simulator, this simulator has been shown to be superior to other simulation algorithms in terms of RMS error by Hajivassiliou, McFadden, and Ruud (1996).

[9] This recursive scheme is also called a *Gibbs cycle* (Geweke 1991).

[10] Since optimizers in software packages such as Gauss, SAS, and others. are *minimizers*, maximum likelihood estimation entails minimization of the negative log-likelihood function. In such a case, the Hessian matrix must be *positive-definite*.

[11] The discussion of the GEV and Nested Logit Models in this section are borrowed from Ben-Akiva and Lerman (1985).

[12] The MNL model, by ignoring attribute-specific random components, that is, ε_{iat} and ε_{ibt}, restricts σ_a^2 and σ_b^2 to be zero.

Bibliography

Batsell, R.R. and Polking, J.C. (1985) A new class of market share models. *Marketing Science*, **4** (3), 177–198.

Ben-Akiva, M. and Lerman, S. (1985) *Discrete Choice Analysis*, MIT Press, Cambridge.

Borsch-Supan, A. and Hajivassiliou, V.A. (1993) Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models. *Journal of Econometrics*, **58**, 347–368.

Chintagunta, P.K. (1992) Estimating a multinomial probit model using the method of simulated moments: a panel data application. *Marketing Science*, **11** (4), 386–407.

Clark, C. (1961) The greatest of a finite set of random variables. *Operations Research*, **9** (3), 145–162.

Debreu, G. (1960) A review of individual choice behavior: a theoretical analysis. *American Economic Review*, **50** (4), 186–188.

Efron, B. (1978) Regression and ANOVA with zero-one data: measures of residual variation. *Journal of the American Statistical Association*, **73**, 113–121.

Geweke, J. (1991) Efficient Simulation From the Multivariate Normal and Student-t Distributions Subject to Linear Constraints, in *Computing Science and Statistics: Proceedings of the Twenty-Third Symposium on the Interface* (ed. E.M. Keramidas), Fairfax: Interface Foundation of North America, Inc., 571–578

Guadagni, P. and Little, J.D.C. (1983) A logit model of brand choice calibrated on scanner data. *Marketing Science*, **2** (3), 203–238.

Hajivassiliou, V.A., McFadden, D.L. and Ruud, P. (1996) Simulation of multivariate normal rectangle probabilities and their derivatives: theoretical and computational results. *Journal of Econometrics*, **72**, 85–134.

Hausman, J. and Wise, D.A. (1978) A conditional probit model for qualitative choice: discrete decisions recognizing interdependence and heterogeneous preferences. *Econometrica*, **46**, 403–426.

Horowitz, J. and Louviere, J. (1993) Testing predicted choices against observations in probabilistic discrete choice models. *Marketing Science*, **12** (3), 270–279.

Jones, J.M. and Zufryden, F.S. (1982) An approach for assessing demographic and price influences on brand purchase behavior. *Journal of Marketing*, **46** (1), 36–46.

Kannan, P.K. and Wright, G. (1991) Modeling and testing structured markets: a nested logit approach. *Marketing Science*, **10** (1), 58–82.

Lerman, S. and Manski, C. (1981) On the use of simulated frequencies to approximate choice probabilities, in *Structural Analysis of Discrete Data with Econometric Applications* (eds C. Manski and D. McFadden), MIT Press, Cambridge, pp. 305–319.

Luce, R. (1959) *Individual Choice Behavior*, John Wiley & Sons, Inc., New York.

McFadden, D. (1974) Conditional logit analysis of qualitative choice behavior, in *Frontiers in Econometrics* (ed. P. Zarembka), Academic Press, New York.

McFadden, D. (1978) Modeling the choice of residential location, in *Spatial Interaction Theory and Residential Location* (ed. A. Karlquist) North Holland, Amsterdam, pp. 75–96.

McFadden, D., Train, K. and Tye, W.B. (1977) An application of diagnostic tests for the independence of irrelevant alternatives property of the multinomial logit model. *Transportation Research Record*, **637**, 39–45.

Papatla, P. and Krishnamurthi, L. (1992) A probit model of choice dynamics. *Marketing Science*, **11** (3), 189–206.

Roberts, J.H. and Lattin, J.M. (1991) A development and testing of a model of consideration set composition. *Journal of Marketing Research*, **28** (4), 429–440.

Swait, J. and Louviere, J. (1993) The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of Marketing Research*, **30** (3), 305–314.

Thurstone, L. (1927) A law of comparative judgment. *Psychological Review*, **34**, 273–286.

Tversky, A. (1972) Elimination by aspects: a theory of choice. *Psychological Review*, **79**, 281–299.

opportunities and challenges in social marketing

Alan R. Andreasen

Brief History

Social marketing has two points of origin, one in the practical world and one in academia. The first instance of application of marketing concepts and tools in a social context is found in family planning programs in India in the late 1960s. Despite significant inputs of international aid, families in underdeveloped and developing countries were becoming worse off because family size growth was outpacing economic growth. At the time, public health programs were minimally effective in part because of their poor reputations and frequent inaccessibility.

A visiting professor at the Indian Institute of Management at Calcutta (IIMC), Peter King, along with businessmen, IIMC professors, and the Ford Foundation brought out the document "Proposals for Family Planning Promotion: A Marketing Plan" in 1964. The concept was to package and advertise branded, donated products and have them distributed by six cooperating international marketers including Hindustan Lever, Lipton, and Union Carbide. The Nirodh program became the world's largest social marketing effort and remained so until the mid-1990s.

In the same period as the Nirodh start-up, the marketing faculty at Northwestern University fomented a sea change in the academic view of marketing by proposing that it be broadened beyond its commercial confines, arguing that "every organization performs marketing-like activities whether or not they are recognized as such ... marketing is a pervasive activity that goes considerably beyond the selling of toothpaste, soap and steel." Two key members of the Northwestern faculty, Philip Kotler and Gerald Zaltman, focused on a particular set of applications which they called for the first time *social marketing* (Kotler and Zaltman, 1971).

Initial academic opposition to the broadening concept (marketing implies "markets") quickly retreated and a number of academics in the United States, many of whom were influenced by the social unrest provoked by the Vietnam war, began to investigate and write about social marketing applications. At the same time, select advertising and public relations organizations such as Porter Novelli began to take on clients with social challenges and existing consulting organizations such as The Futures Group, Population Services International, and the Academy for Educational Development also secured social marketing contracts from government agencies such as USAID (United States Agency for International Development), CDC (Centers for Disease Control and Prevention), and the World Bank.

The growth of interest in social marketing on the scholarly and practical sides approximated a typical diffusion process which is still in the growth phase. Research and applications grew relatively slowly at first because, on the academic side, articles of top journal quality were slow in appearing and, on the practical side, because clients were often unfamiliar with or confused about the concept. Both impediments have largely been overcome.

There are now at least eight textbooks and various specialized volumes. The field has its own journal, the *Social Marketing Quarterly*, and social marketing articles routinely appear in the *Journal of Consumer Research*, *Journal of Public Policy and Marketing*, and other major journals around the world. There are reading books on social marketing ethics, annual conferences, and special sessions within other mainstream conferences that are now devoted to social marketing. Chapters on social marketing appear in books in nonmarketing disciplines such as health promotion and health communication. Social marketing as a strategic platform has been introduced into US agencies like the Centers for Disease Control and Prevention and into a range of UK programs of the National Health Service.

Definitional Challenges

One of the early – and in one respect continuing – problems for social marketing was confusion about what the term meant and how it differed from nonprofit marketing, socially responsible marketing, social advertising and, most recently, social network marketing. In the earliest period, social marketing was seen as product marketing (as in the family planning work and later in many

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

child survival medical interventions). However, the field slowly evolved to encompass service applications (mammography screening and drug counseling campaigns) and eventually "pure behaviors" (recycling and walking more). Along the way, given the dominance of health applications, it faced challenges to differentiate itself from health education or health communication. A further of source of confusion was scholars and promoters who claimed that it encompassed the marketing of *ideas*, which led critics to argue that it was just a fancy version of advocacy or propaganda. Many asked: "Isn't this what the Advertising Council has been doing since World War II?"

By the mid-1990s, a consensus emerged around the fundamental notion that the objective was to influence *behavior* and not just attitudes or knowledge. It was argued that the commercial marketers whose concepts and tools social marketers apply do not settle for program impacts solely on knowledge and attitudes. If marketing efforts do not result in sales, marketing managers or advertising agencies are not rewarded and campaigns – even products and brands – are scrapped. This author's own definition is that social marketing is the application of commercial marketing concepts and tools to influence the voluntary behavior of target audiences to improve their lives or the society of which they are a part. It is made clear that the goal is to influence *problem* behaviors; not simply behaviors like voting, charitable giving, or attending the arts. The latter fall into the realm of nonprofit marketing more broadly construed.

The focus on behavioral outcomes also distinguishes it from social advertising which has communication goals which may be needed to provoke or assist a behavioral outcome. It is also distinguishable from social network marketing which also can be a tool of social marketing but is not equivalent to it.

Two more definitional challenges revolve around social marketing strategy rather than tactics. The first is whether social marketing can be used to promote racial discrimination or Aryan superiority. The author's own view is that social marketers are behavior influence specialists – as are marketers in the private sector. The choice of objectives and overall goals of some strategy is not their province – although they may provide advice on feasibility. Social marketers need to make ethical choices about the strategies to which they lend their talents – as do commercial marketers asked to promote cigarette brands.

The second strategic issue is whether social marketing should focus solely on individuals or groups exhibiting – or likely to exhibit – problem behaviors, what is often referred to as *downstream* targets. Michael Rothschild (1999) proposed that social marketing scholars and researchers pay more attention to the distinction made by MacInnis, Moorman, and Jaworski (1991) that behaviors – both commercial and social – require three preconditions – the *motivation* to act, the *opportunity* to act, and the *ability* to act (often denoted by the acronym MOA). The obesity crisis has provided a particularly compelling example of the need to focus strategically not just on overweight kids and their families but also on what social marketers call *upstream* targets – school administrators who can change cafeteria menus, banish soft drink machines, and add physical education classes, or city council members who can provide playgrounds or increase police protection in areas where obese kids might play.

In his most recent book Andreasen, (2007), the author has made the case that social marketing has a proper role to play in influencing upstream targets as well as those who exhibit – or might exhibit – problem behaviors. This point deserves expansion here.

"Consumer" Behavior

First, a conceptually important distinction in social marketing is that its definition focuses on "target audiences" and not "consumers." The distinction was made early on, in part, to make the application of marketing concepts and tools more palatable to those in the social sector who considered marketing as somehow evil. However, it also reflects the fact that the objective of many social marketing campaigns is not "consumption." The focus is behavior. In the private sector, the behavior in question is acquisition and consumption of specific goods or services by *consumers*. In social campaigns, such as those directed at exercise or recycling, it is

confusing – even off-putting – to imply that the goal is to influence some sort of "consumption." As a consequence, this author has advocated the more general – and more encompassing – term of *target audience behavior*, a term that subsumes both commercial and social objectives.

This is not merely a semantic challenge. It may be argued that marketing is all about influencing behavior. Creating intellectual silos, wherein social marketing is seen as somehow different and a special (and minor) area of application of marketing concepts and tools, is a taxonomic construct that inhibits the opportunity to create a more generic view of the field. That is, we should ourselves, as behavior influence researchers, seek commonalities across behavioral contexts. For example, rather than dividing the field into purchasing and social behaviors, would we not advance the field further if we ask more generic questions about behavioral influence? For example, how does achieving the goal of *starting* a behavior differ from *continuing* a behavior or *switching* a behavior? What do we know about stopping a behavior? Is preventing someone from engaging in a behavior (stealing videos, using drugs, engaging in risky sex) different from getting them to start a behavior? Are there general principles that apply across sectors in starting or switching behaviors?

Further, if influencing behavior is our strength, then can we consider efforts to influence upstream individuals a legitimate subject of interest? Is getting a city council member to vote "yes" to fund a school playground in a poor neighborhood an example of influencing target audience behavior? What of cafeteria managers, editorial directors at TV stations, or the person who decides what information goes on menu boards at a Burger King? The author's argument is that the answer is yes and, further, many of these situations would present opportunities to study *stopping* behavior in more depth.

There are already a number of commonalities in the way that marketing managers in both sectors approach their behavioral challenges. The commercial concept of value creation, for example, is equally relevant in both sectors. A social marketer seeks to offer target audiences compelling value propositions comprising attractive benefits and minimal costs so as to provoke a desired behavioral outcome in which benefits are exchanged for costs (Bagozzi, 1978; Bagozzi and Warshaw, 1990). Thus, the truth® campaign of the American Legacy Foundation has the goal of inducing teenagers not to take up smoking (a behavioral outcome). The campaign offers the target audience *benefits* such as a sense of pride in not being manipulated by the tobacco industry and a feeling of comradeship in being part of an "anti-tobacco movement of teens." These benefits have been found to more than compensate for the costs of not having specific (smoker) friends or enduring the taunts of teen smokers who implicitly argue that smoking is "cool."

The future of "consumer behavior" research should include special issues of major journals and conference sessions on switching, stopping, and starting behaviors along with those on brand preference, cognitive processing, and customer satisfaction. Further, one would hope that scholars interested in the latter topics would, from time to time, ask whether and how their theories and findings would apply in different realms, not just in the commercial marketplace.

Target Audience Behavioral Research

Social marketing comprises an area of application of marketing concepts and tools similar to B2B marketing and services marketing. Textbooks and strategic plans therefore insist on thorough target audience research in planning interventions, pretesting of such interventions, and regularly tracking progress.[1] Social marketing applications also urge sophisticated segmentation approaches and close attention to positioning and branding possibilities.

The research in the area of social marketing divides into three broad categories. First, there are macrolevel evaluation studies that seek to ascertain whether a particular intervention yielded desired effects and, where possible, which elements of the specific campaign seemed to have been most impactful. Second, there are midlevel studies of particular tactics, typically variations on communication approaches. Third, there are microlevel studies that seek to discover mental processing pathways between inputs and desired effects.

Typical of macrolevel research is an evaluation of 54 social marketing programs carried out

in 2006 by Stead *et al.* (2007). The authors' conclusions are characteristic:

> A majority of the interventions, which sought to prevent youth smoking, alcohol use and illicit drug, use reported significant positive effects in the short term. Effects tended to dissipate in the medium and longer term, although several of the tobacco and alcohol interventions still displayed some positive effects two years after the intervention. These results are broadly comparable with systematic reviews of other types of substance use prevention interventions . . . . The evidence is more mixed for adult smoking cessation, although small numbers of programmes were nonetheless effective in this area. There is modest evidence of impact on levels of physical activity and psychosocial outcomes, with an apparently weaker effect on physical activity related physiological outcomes.

Two well-funded campaigns in the United States that represent social marketing success stories are the VERB campaign of the Centers for Disease Control and Prevention and the truth® campaign of the American Legacy Foundation. The VERB campaign used sophisticated branding and media approaches focused on children 9–13 years as well as upstream influencers. It used "commercial marketing methods to advertise being physically active as cool, fun, and a chance to have a good time with friends." Baseline and 1-year follow up surveys of 3120 parent–child dyads were conducted in 2002 and 2003. Huhman *et al.* (2005) concluded that "The VERB campaign achieved high levels of awareness in 1 year." They also noted particularly marked effects on the behavior of younger children, girls, children whose parents had less than a high school education, children from urban areas that were densely populated, and children who were "low active at baseline." (Ironically, after spending $339 million over 5 years, funding for this successful campaign was withdrawn in 2006).

The truth® campaign has a much longer history. It began under the Florida State Government with funding from one of four early settlements of state suits against the tobacco industry and migrated to the American Legacy Foundation which was created with funding from the subsequent master settlement negotiated with 48 states and the District of Columbia. The campaign achieved dramatic effects on smoking initiation in Florida, results that are paralleled in research on the national campaign (Farrelly *et al.* 2002).

Midlevel and microlevel social marketing research studies both exhibit a key difficulty in modeling and understanding social marketing effects at the micro and campaign level. Despite an earlier argument that marketing is all about behavior, whatever the sector, it is the case that behaviors to be influenced in the social sector often have one of three characteristics:

1. There is no observable behavior that one can measure that would mark behavioral outcomes – for example, not smoking and not using drugs.
2. The objective is to instill a future pattern of repeated behaviors – for example, exercising regularly, advancing girls' education and employment in developing countries, regularly taking one's medicine, and recycling.
3. The major benefit constitutes an expectation (promise?) of something that takes place many years in the future – for example, a longer life – or that comprises a "nonoutcome" such as a stroke or breast cancer both of which may be perceived as probabilities of such outcomes.

As a consequence, considerable focus in social marketing research has been on measuring and tracking *intentions* to behave. Fortunately, these intentions have been a key measure in commercial consumer behavior research especially in microlevel consumer behavior experiments. It has also been a particular focus in the area of health marketing where perhaps 80–85% of social marketing applications have been initiated. The latter has the important advantage for those eager to conduct target audience research in social marketing that there exists a significant body of theory and research, much of it seeking to explain or predict socially important behavior from measures of intentions.

A useful summary of much of this work is provided in the comprehensive 2006 meta-analysis by Webb and Sheeran (2006) of 47 experimental studies of the linkage between intentions and behavior change.[2] The models

that undergirded many of these studies will be familiar to traditional marketing scholars. Eighty-one percent were grounded in one of four models. The most common (29%) was Fishbein and Ajzen's Theory of Reasoned Action (TRA) or the modification in Ajzen's Theory of Planned Behavior (TPB). This was followed by Bandura's social-cognitive theory (21%), Ronald Rogers' Protection Motivation Theory (18%), and Hochbaum and Rosenstock's Health Belief Model (13%). As most consumer behavior researchers know, the TRA/TPB models specify that intentions and ultimately behaviors are driven by three factors: (i) beliefs about outcomes and the evaluation of those outcomes; (ii) normative beliefs about the wishes of significant others and the motivation to comply with those wishes; and (iii) self-efficacy – "control beliefs and perceived power." Other factors that can influence these three central components as well as factors that potentially impact the link between intentions and behavior are summarized by Fishbein and Cappella (2006). The potential modifiers between intentions and behavior are environmental factors and skills and abilities. Six sets of background factors can influence behavioral, normative, and control beliefs. These are

- past behavior;
- demographics and culture;
- attitudes toward targets (stereotypes and stigma);
- personality, moods, and emotions;
- other individual difference variables (perceived risk);
- intervention exposure and media exposure.

Bandura's model focuses primarily on self-efficacy beliefs – one's ability to act and to secure a specific outcome. Roger's Protection Motivation Model emphasizes individual perceptions of vulnerability to health threats and appraisals of their coping skills – the perception of the efficacy and costs of recommended responses, The Health Belief Model is similar in that it assesses individual's perceptions of health threats and their ability to favorably respond to them. These models are often used by social marketing researchers as in Pechmann *et al.*'s use of the protection motivation model to assess the effectiveness of antismoking advertisements on adolescents (Pechmann *et al.*, 2003).

All of these models focus on factors influencing intentions. Webb and Sheeran (2006) particularly focused on the extent to which intentions actually led to behavior. Three sets of factors were considered: the theoretical base and the methods and perceived sources of the intervention. Of potential future research interest was the finding that intentions were less predictive of behavior under two conditions – prior habit formation and social reactions to the particular health behavior.

Social reactions have been of particular interest to Robert Cialdini and colleagues. Cialdini found that perceived social norms have a strong effect on actual behaviors. Ironically, he also found that subjects in their studies perceived norms to be much less influential than they actually were (Kallgren, Reno, and Cialdini, 2000; Nolan *et al.*, 2008).

Many approaches to socially important behaviors have considered intentions as a variable represented as probabilities of action. To augment this approach, Prochaska and DiClemente introduced what they called the *Transtheoretical Model of Behavior*. The focus of this approach is that, for the high involvement kinds of behaviors that are the focus of much social marketing research, it is more reasonable to think of target individuals proceeding not in a single step from considering a behavior to acting on it (or not). Rather, these researchers proposed that target audiences proceed through six "stages of change." These are as follows:

- Precontemplation – where the target audiences are either unaware of the need for the behavior or consider it not of relevance for them.
- Contemplation – where they are actively considering the action – a stage that can take considerable time and which has been thought of as "early" and "late" contemplation.
- Preparation – where the individuals have decided to go forward and are, perhaps, assembling the necessary skills, social support, and the like to accomplish the behavior.

- Action – where the individuals are undertaking their initial attempts at behavior change (which may be all that is needed, for example, securing an important inoculation).
- Maintenance – where the individuals take up a new pattern of behavior and, preferably, make it a habit.
- Termination – either abandoning or completing a course of action.

Their model has been applied with considerable success to such behaviors as smoking, drinking, condom use, responsible eating, and exercise (Prochaska, DiClemente, and Norcross, 1992; Prochaska and Velicer, 1997).

Recent Developments

The twenty-first century has seen a number of significant developments that raise important opportunities for research. First, as noted earlier, social marketers have begun to pay more attention to environmental factors that often provide significant social influence. Social norms (already cited) are one set of factors. Social norms, of course, are themselves not fixed. They both evolve over time and are themselves potentially capable of being influenced. In this regard, the author has suggested elsewhere that social marketers investigate the formation and influence of social agendas. Important precursors of social influence may be the relative importance of the behavior on the general public agenda. Such agendas ebb and flow in particular stages and are very often influenced by media agendas.

A second recent focus that reflects attention to upstream factors is the attention to behavioral economics. This line of thinking stems from conservative thinkers who are reluctant to promote strong campaigns to change behaviors but, instead, argue for changes in environmental conditions that can "nudge" people toward doing the right thing (Thaler and Sunstein, 2008). Wansink's series of studies on such influences on overeating suggest possible courses of research along these lines (Wansink, 2004).

A third development has been an expansion of the areas of interest beyond health care. As environmental issues rise on political or public agendas, more social marketers are exploring intervention alternatives. McKenzie-Mohr and Smith (1999) have emphasized that such issues call for much more attention to community involvement in solutions. In a different realm, Lusardi, Keller, and Keller (2009) have used social marketing concepts to study financial decision making.

A fourth important development is the merging of social marketing interests with those of social entrepreneurs including corporations. In recent years, private sector businesses have increasingly become interested in bringing about social change (Hess, Rogovsky, and Dunfee, 2002). They do so, not because they are generous and big – hearted – although they may be – but because they envision potential strategic payoffs to their own bottom lines. For example, Hindustan Lever has found that it could create a longer lasting, smaller bar of Lifebouy soap, thereby combating sanitation challenges and expanding the total market. Coca Cola found that it needed to bring its marketing and promotional skills to bear on the HIV/AIDS crisis in Africa where it found many of its workers afflicted with the disease.

Future Research Needs

As several leading scholars with interests in social marketing have said, the potential for innovative research in noncommercial settings is significant. Part of this is due to several characteristics of the behaviors that social marketers are attempting to influence. Among these are behaviors where

1. there are no concrete products or services involved – for example, exercise;
2. they are really *non* behaviors – for example, not smoking, not engaging in violence against spouses or children, and not using drugs;
3. they promise unobservable benefits (without instrumentation) – for example, lowered blood pressure;
4. they promise benefits long in the future – for example, longer life from exercise or better diets and a happier retirement from today's savings;
5. they often require significant trials before success is achieved – for example, stopping smoking or losing weight;

6. the behaviors are heavily influenced by significant others – for example, health practices of young African women under the watchful eyes of their mothers-in-law;
7. the benefits of the behavior accrue to others while the costs are personal – for example, recycling;
8. interventions are constrained by public scrutiny or by social norms – for example, HIV/AIDS campaigns in many countries and family planning initiatives by the US government (especially under conservative administrations).

There is also an intriguing challenge that, to the author's knowledge, has not been addressed in consumer behavior studies in commercial settings. This is the fact that, in social marketing campaigns, there are often multiple *types* of target audiences that will be affected by a particular campaign or tactic but may not respond in the same way. Resources for social marketing campaigns come from five principal sources: revenues from clients (e.g., Goodwill clothing sales), funding by foundations, government contracts, individual donations, volunteer commitments, and various forms of involvement by corporations. Typical consumer behavior research assesses responses – or likely responses – for a single type of target – consumers. In nonprofit contexts, it may be that volunteers or donors will respond differently (and perhaps negatively) to tactics that are positively effective with donors or foundations. What if alternative AIDS campaign messages yield different responses from at-risk populations, volunteers, and corporate partners? How does one choose among them?

Branding is potentially highly valuable for social marketing programs. It has been used to brand specific products (e.g., Nirodh condoms) or services (Family of the Future health clinics). However, there are two problems for campaigns that do not involve products or services. What, for example, should branding's role be in tackling the obesity problem? Is there "customer value" to be derived from branding an intervention to get someone eating better or exercising more? Will there be more frequent or more lasting outcomes if a target audience member follows the recommended behaviors of a branded social marketing campaign? Are slogans – *Don't Mess with Texas* or *Click It or Ticket* – equivalent of brands?

Another challenge alluded to above is the fact that many desirable social outcomes require that the target audience adopt a different behavioral pattern – often over a lifetime – whether it is a better diet, regular exercise, condom use, recycling, and so on. Experience shows that lifestyle changes often suffer setbacks and reversals. Sometimes, it is because the behavior seems difficult (regular exercise) or because outcomes are not obvious on a daily basis (lowered blood pressure). How does one study cognitive processes yielding not just a one-time behavior but a pattern of future behaviors? Useful beginnings are found in Bagozzi and Warshaw's, 1990 "theory of trying." More recent research by Kahn and Luce (2003) may add to this by looking at the role of feedback on continued behavior, in their case, the impact of "false positives" from mammographies on repeated test-taking.

Another research need is to pay more attention to intervention strategies other than communications and message alternatives. It is understandable that those describing and investigating social marketing interventions would focus on communications. Clever and unconventional advertising approaches are seductive as a means of describing and bragging about campaign performance. (Indeed, the author's colleagues have observed that his own writings often exhibit this tendency.) There are three fundamental problems with this distorted emphasis.

First, it offers ammunition to those who argue that we bring nothing to problematic areas that are already dominated by communications specialist. Second, it sorely underutilizes marketing's robust armamentarium – for example, other elements of the "4Ps." In the early family planning programs, a critical contributor to success was the establishment of far flung and *reliable* distribution networks for family planning products. This was essential to a poor mother seeking to limit her family size because she could rely on the availability of a new cycle of oral pills when she had enough money to buy another round. Equally important was the pricing of family planning products. Research has shown that free products are perceived as of lower quality and reliability than

products (and services) that cost something, but how much? Rules of thumb were used initially but this is one of those realms where careful target audience research by marketing scholars could not only help organizations make tactical decisions about pricing but also broaden our understanding and interpretation of pricing signals by the poor and not-so-poor. Such research would also have the virtue of fueling arguments against social critics who rail against social marketing approaches that charge anything to the poor as immoral.

Third, it reinforces the myopic view that all that is needed for the desired behavioral influence that the target audience be sufficiently motivated. However, as Rothschild and others keep emphasizing and the cognitive models of Fishbein and others make clear, social behaviors also require that target audiences have the ability and opportunity to act and keep acting. This requires social marketing scholars to pay greater attention to influencing the behaviors of upstream target audiences, not typically considered to be "consumers" in the way the term is construed in traditional texts. That is, for a society to have major impact on childhood obesity, school administrators have to provide better lunches and breakfasts and safer school playgrounds. Similarly, McDonald's marketers need to bring marketing skills to the problem and legislators (sometimes) have to pass laws requiring better food labeling, bans on transfats and/or funding for new parks and bike paths.

If consumer behavior research is about influencing behaviors to increase desirable organizational outcomes (more profits, more donations, or more grants) and if we extend our consideration to the task of improving individual and societal welfare, then can we not consider a school administrator an important target audience whose behavior is needed to yield better social outcomes? If so – as I believe it is – then this opens up a dramatic new array of opportunities to study a whole new class of targets and potentially advance our understanding of how marketing's 4Ps can influence target audiences to protect the poor from poverty and disease, improve the world's physical environments, and improve the health and financial well-being of all.

Endnotes

[1] This contrasts with historical social science approaches to project research that emphasized benchmarking and project-end evaluations.

[2] The 47 studies were selected out of a broader universe of 221 studies in that both intentions and behavior were measured and that the latter measures were separated in time. That is, intentions were measured and behavior was tracked at some later point.

Bibliography

Andreasen, A.R. (2007) *Social Marketing in the 21st Century*, Sage, Thousand Oaks.

Bagozzi, R.P. (1978) Marketing as exchange: a theory of transactions in the marketplace. *American Behavioral Scientist*, **21**, 535–556.

Bagozzi, R.P. and Warshaw, P.R. (1990) Trying to consume. *Journal of Consumer Research*, **17**, 127–140.

Farrelly, M.C., Healton, C., Davis, K.C. *et al.* (2002) Getting to the truth: evaluating national tobacco countermarketing campaigns. *American Journal of Public Health*, **92**, 901–907.

Fishbein, M. and Cappella, J.N. (2006) The role of theory in developing effective health communications. *Journal of Communications*, **56**, S1–S17.

Hess, D., Rogovsky, N., and Dunfee, T.W. (2002) The next wave of corporate community involvement. *California Management Review*, **44** (2), 110–125.

Huhman, M., Potter, L.D., Wong, F.L. *et al.* (2005) Effects of a mass media campaign to increase physical activity among children: year-1 results of the VERB campaign. *Pediatrics*, **116**, 277–284.

Kahn, B.E. and Luce, M.F. (2003) Understanding high-stakes consumer decisions: mammography adherence following false alarm test results. *Marketing Science*, **22** (3), 393–410.

Kallgren, C.A., Reno, R.R., and Cialdini, R.B. (2000) A focus theory of normative conduct: when norms do and do not affect behavior. *Personality and Social Psychology Bulletin*, **26**, 1002–1012.

Kotler, P. and Zaltman, G. (1971) Social marketing: an approach to planned social change. *Journal of Marketing*, **35**, 3–12.

Lusardi, A., Keller, P.A., and Keller, A. (2009) New ways to make people save: a social marketing approach, in *Overcoming the Saving Slump: How to Increase the Effectiveness of Financial Education and Saving Programs* (ed. A. Lusardi), University of Chicago Press, Chicago.

MacInnis, D.J., Moorman, C., and Jaworski, B.J. (1991) Enhancing and measuring consumers' motivation, opportunity and ability to process brand information from ads. *Journal of Marketing*, **55**, 32–53.

McKenzie-Mohr, D. and Smith, W. (1999) *Fostering Sustainable Behavior: An Introduction to Community-Based Social Marketing*, New Society Publications, New York.

Michie, S., Johnson, M., Francis, J. *et al*. (2008) From theory to intervention: mapping theoretically derived behavioural determinants to behavior change techniques. *Applied Psychology: An International Review*, **57** (4), 660–680.

Nolan, J.M., Schultz, P.W., Cialdini, R.B. *et al*. (2008) Normative social influence is underdetected. *Personality and Social Psychology Bulletin*, **34**, 913–923.

Pechmann, C., Zhao, G., Goldberg, M.E., and Reibling, E.T. (2003) What to convey in antismoking advertisements for adolescents: the use of protection motivation theory to identify effective message themes. *Journal of Marketing*, **67**, 1–18.

Prahalad, C.K. (2004) *The Fortune at the Bottom of the Pyramid: Eradicating Poverty Through Profits*, Wharton School Publishers, Philadelphia.

Prochaska, J.O., DiClemente, C.C., and Norcross, J.C. (1992) In search of how people change: applications to addictive behaviors. *American Psychologist*, **47** (9), 1102–1114.

Prochaska, J.O. and Velicer, W.F. (1997) The transtheoretical model of health behavior change. *American Journal of Health Promotion*, **12**, 38–48.

Rothschild, M.L. (1999) Carrot sticks, and promises: a conceptual framework for the management of public health and social issues behavior. *Journal of Marketing*, **63**, 24–37.

Sly, D.F., Hopkins, R.S., Trapido, Ed., and Ray, S. (2001) Influence of a counteradvertising media campaign on initiation of smoking: the Florida "truth" campaign. *American Journal of Public Health*, **91** (2), 233–238.

Thaler, R.H. and Sunstein, C.R. (2008) *Nudge: Improving Decisions about Health, Wealth, and Happiness*, Yale University Press, New Haven.

Stead, M., Gordon, R., Angus, K., and McDermott, L. (2007) A systematic review of social marketing effectiveness. *Health Education*, **107** (2), 126–191.

Webb, T.L. and Sheeran, P. (2006) Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychological Bulletin*, **132** (2), 249–268.

Wansink, B. (2004) Environmental factors that increase the food intake and consumption volume of unknowing consumers. *Annual Review of Nutrition*, **24**, 454–479.

brand community

Utpal M. Dholakia and René Algesheimer

Introduction

Much of consumer behavior is social, impelled by social motives and joint decision-making processes, shaped by socially constructed and shared information, and governed by social norms, influences, rituals, traditions, and taboos. Over the last decade, consumer researchers have increasingly turned their attention to social processes. Among other things, they have studied family and couples decision making, social influences on consumer decision making and self-regulation, consumer subcultures, and the effects of community on consumer choices, attitudes, and behaviors.

A social concept that has emerged as important in explaining and understanding consumer behavior is that of a consumer collective called a *brand community*. The concept of brand community is powerful because it reflects both social process and cultural meaning, stems from a number of converging environmental trends such as the ascendance of the Internet, the decline of traditional family and community, and the growing power of individual consumers in marketplace transactions, and influences consequential marketing outcomes. In this article, we elaborate on this concept by providing a definition of brand community, introduce and elaborate on different types of brand communities, and examine the consequences of consumer participation in them.

Definition of a Brand Community

Brand communities serve many different functions for consumers and firms. For consumers, brand communities act as conduits of information, channels for solving product-related problems and learning how to use its features, places for finding new friends for social support and for meeting existing friends, and as a means for self-expression through creation and sharing of symbolic content. For firms, brand communities are low-cost, high-efficacy marketing programs which can achieve a number of different marketing objectives simultaneously: the abilities to conduct quick and low-cost marketing research with the target audience, deliver prompt customer service at low expense, educate and socialize new customers, strengthen attachment to the firm's brand for existing customers, and increase the frequency and loyalty of customer purchase behaviors.

We define a *brand community* as *a collective of consumers organized around one particular brand, which is sustained through repeated online and/or offline social interactions and communication among its members who possess a consciousness of kind, feel moral responsibility toward one another, and embrace and propagate the collective's rituals and traditions*. Several aspects of this definition are important and worth elaborating on.

Role of brand in brand community. First, a brand community is a social collective organized around one particular brand, which means that the collective comprises of consumers who have at least some heightened enduring interest in that brand. Such an interest may stem from an attachment to the brand itself, from a more general interest in the product category to which the brand belongs, or (more likely) both. In the prototypical example of a brand community involving Harley Davidson Motorcycle owners, the so-called HOG (Harley Owners Group), many members have a fanatical devotion to the Harley Davidson brand, and view the biker lifestyle symbolized by the brand in idealistic terms. Generally speaking, brand community members are among the most ardent enthusiasts of the brand. Consumer researchers have found brand communities comprised of enthusiasts to exist across a range of product categories, from highly complex technological products like enterprise software (e.g., Oracle and Hewlett Packard) and construction equipment (e.g., John Deere), to consumer durables like virtually all brands of cars, motorcycles, and video game consoles, to low-cost consumable food brands like the Nutella chocolate hazelnut spread, Coca-Cola, and Starburst candy.

Role of communication in brand community. The second part of the definition explains the process by which the brand community sustains itself. Regular social interactions and communication

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

between members, accomplished through online channels such as bulletin boards, chat-rooms, and email lists, and via offline means such as face-to-face meetings, events, and gatherings, is essential not only for the community's business to be conducted but also, perhaps more importantly, for the relationships between community members to form and strengthen. It is through timely and convenient communication ability that the brand community is able to solve the problems of individual members, deliver technical service and support to them, generate consumer feedback and new product ideas, and deepen participants' knowledge of one another and the strength of their relationships. For individual participants, communicating with others within the brand community serves specific functional purposes such as solving a particular product-related problem or learning how to use a product feature, or may simply provide the means to have a pleasant and enjoyable communal experience.

The essential markers of brand community. The remaining part of our brand community definition describes the three markers that sociologists stipulate as essential for any social collective to be truly considered a community: (i) a consciousness of kind, (ii) a sense of moral responsibility, and (iii) the knowledge and acceptance of the collective's rituals and traditions.

The first core marker of community, the so-called *consciousness of kind*, refers to the intrinsic connection that community members feel toward one another through a sense of belonging to the group, and a sense of difference or separation from those who are nonmembers. Consciousness of kind is the force driving the cognitive categorization of in-groups and out-groups and biased behavior favoring the in-groups by consumers, and a primary reason why brand communities are such effective marketing programs. Compared to advertising or direct marketing, brand communities muster the consumers' intrinsic interests and motivations in support of the brand, and in opposition to competing brands. For example, Harley Davidson riders belonging to a HOG believe that they share similar attributes, values, and views of life, and thus feel intrinsically connected to each other. At the same time, they may view riders of other brands such as those who own Honda or Suzuki motorcycles in a negative light. Such beliefs and feelings strongly influence their behaviors.

As Cova 1997 points out in introducing the notion of "linking value," even if brand community members have never met before in a face-to-face setting (as is often the case with brand community participants), they can still experience the connection and the feeling that they know one another. The common social link shared with other members through their interest in the brand sustains the community, even without actual physical interaction among its members.

The second core community marker is a feeling or sense of moral responsibility or obligation toward the community itself and other brand community members, and may include a concern for their well-being as expressed through acts of help or social support such as by teaching newer members how to use the product, and educating them about the practices and norms of the community. The sense of moral responsibility can extend to the community as a whole, and is evident when members make the effort to enforce communal rules and shared values such as fairness, and go out of their way to recruit new members to the community. Usually, such a sense of responsibility is cultivated over time as a member participates in and comes to identify with the community. It deepens with the consumer's experience and tenure in the brand community.

Muniz and O'Guinn (2001) insightfully observed that brand communities are *communities of limited liability* in the sense that for individual consumers they are intentional, voluntary, and partial in the level of involvement they engender, yet are vital to contemporary life and convey significant meaning to the consumer.

The third core community marker is the knowledge and acceptance of the brand community's shared rituals and traditions by its members. As a social collective matures into a community, it develops various rituals and traditions. In an important way, it is these rituals and traditions, which may include such things as narratives of the brand's origins and history, celebrations, brand stories and myths, and ritualistic utterances and actions, which

create the sense of the in-group and affirm the cohesiveness of brand community members. For example, it is customary for members of eBay's brand community to recount personal acts of devotion to the eBay site in its chat forums, and instances where they converted a friend or a loved one from the use of a competing site such as amazon.com.

In addition to these markers, an important characteristic of the brand community is the dynamics of the collectivity. Rather than being a static, stabilized state of social relations previously defined as a community, brand communities are dynamic and active. As a consequence, community members are neither passive, nor reactive individuals that only behave in accordance to internal forces, for example, compliance or self-esteem, or external forces, for example, group norms, or peer pressure, of the community. Members actively choose their community and decide when to leave it. Furthermore, ongoing interactions between and activities of community members shape and change the appearance and structure of the community itself. For example, how rituals and traditions are celebrated in the community influences the emergence of its cultural symbols. On the other hand, the brand communities' structure, its norms and set of rules, its rituals and traditions influence individual's future activities, social interaction, and consumption patterns. Overall, one can say that brand communities are not only "produced" by their members but also "reproduced" by social interactions between their members that reinforce the community. Thus, brand communities are dynamic social phenomena.

Types of Brand Communities

In further understanding brand communities, it is important to distinguish between (i) brand communities that are comprised of consumer networks and those that consist primarily of small friendship groups and (ii) brand communities that are centrally managed by the firm's managers versus those that are often decentrally organized and managed through grassroots efforts by customer enthusiasts.

Network-based and small-group-based brand communities. In describing communities, sociologists make the distinction between *neighborhood solidarities*, which they define as tightly bound, densely knit groups with strong relationships between members, and *social networks*, which are loosely bound, sparsely knit networks of members sharing weak and narrowly defined relationships with one another. Whereas neighborhood solidarities tend to be geographically conjoint, where each member knows everyone else and relies on them for a wide variety of social support, social networks are usually geographically dispersed groups that interact with one another for a specific reason, without prior planning.

Social psychologists similarly distinguish between *common bond* and *common identity* groups. Whereas the bond between members is the glue holding the group together in common bond groups, the attachment depends on identification to the whole group, in common identity groups. Common bond groups correspond to neighborhood solidarities and common identity groups to social networks. These distinctions, of viewing the community as either more or less the same small group of individuals with each of whom the consumer has relationships, or viewing it as a venue where numerous, dynamically changing people (strangers or acquaintances) with shared interests or goals meet, is useful in classifying brand communities.

In some instances, the consumer thinks of the brand community primarily as a venue, and only superficially associates it with any particular individual(s) within it. For instance, a consumer may log on to the bulletin-board of a software company because he/she has a problem that needs solving. In this case, his/her main interest is in solving the problem; there is no expectation or inclination to meet, chat, or socialize with any particular community member. Likewise, an engaged *yelp.com* or *amazon.com* consumer may read and benefit from reviews offered by others, without personal knowledge of, or relationships with, the reviewers. A brand community defined this way, that is, as "a network of relationships among consumers organized around a shared interest in the brand and promoted mainly via online channels, where intellectual and utilitarian support is primary and emotional

support is secondary" is a "*network-based brand community*."

In other cases, the brand community's member may identify primarily with a specific small group (or groups) of consumers, rather than with the venue in which the community meets. For example, a software developer may log on to the community chat-room specifically to chat with his/her geographically distant buddy group of kindred software developers every week to trade ideas, learn new concepts, and to socialize with them. Here, the developer's focus is on communication with his/her peer group that he/she knows personally, rather than on the brand community venue. Such a brand community, "constituted by individuals with a dense web of relationships and a consciously shared social identity interacting together as a group, in order to accomplish a wider range of jointly conceived and held goals, to express mutual sentiments and commitments, and to maintain existing relationships," is a "*small-group-based brand community*."

Differences between network-based and small-group-based brand communities. There are several important differences between network-based and small-group-based brand communities. First, not surprisingly, the specific group with which the consumer interacts holds greater importance for members of small-group-based when compared to network-based brand communities. This is because the individual knows everyone else personally, and in many cases, may have shared histories and close personal relationships with them. As a result, relationships between community members are stronger, more resilient, and more stable than those in network-based brand communities, where members participate primarily to achieve functional goals (e.g., to trouble-shoot a problem) and have tenuous, short-lived, and easily severed ties.

Accentuating the group's importance for small-group-based brand community members is also why the brand community venue is often only one of a number of places where such groups meet. Online social interactions are supplemented by face-to-face and other offline forms of interactions. For instance, a small group of HOG members may not only chat online with one another periodically in the course of a week but also meet on weekdays for coffee and fellowship, and on weekends for group outings. In contrast, network-based community members are more likely to interact with each other exclusively through the brand community venue.

The two brand communities also differ in the range of activities its members engage in. Network-based community members are likely to engage primarily in narrow, instrumental, brand-related activities. In contrast, small-group-based community members engage in broad-based activities. For Harley riders belonging to small group brand communities, for example, social interactions occur through group rides for purely recreational purposes or for more formal goals such as fundraising (e.g., a rally for raising funds for the victims of an earthquake or a hurricane), competitions, political protests (e.g., anti-helmet law rallies), or community service. Even more frequently, small group members come together to meet at a pub or restaurant or to mutually examine the latest bikes and accessories at a nearby dealership.

These differences create what we call the *loyalty-influence paradox* in brand communities. Customers belonging to small-group-based brand communities are less likely to be loyal to a particular venue offered to them by the firm to interact in than members of network-based brand communities; yet, at the same time, small-group-based brand community members are more likely to be influenced by the social interactions with other members in the venue when they do participate than those belonging to network-based brand communities. Thus, loyalty to the brand community is inversely correlated with its influence on its members.

The distinctions between network-based and small-group-based brand communities also have implications for managers of these communities. Whereas the primary managerial objective in a network-based brand community is to match individual motives, for example, to find members willing to help solve product-related problems of those who have the questions, in the case of small-group-based brand communities, the main goal is to satisfy the motivations of group members to socialize

with one another by providing them various applications that facilitate social interactions. Thus, tools and applications such as buddy lists, instant messaging, providing status updates, and sharing of personal history are likely to be more valued in small-group-based brand communities. In contrast, applications that allow specific functional goals to be reached such as an archive of product-related problem solutions, an "ask-an-expert" service, and a reputation system which rewards problem-solving of other members, are more useful in network-based brand communities. The differences also mean that the marketers' role is starkly different in the two cases: to be active information providers and problem solvers in the case of network-based brand communities, but to be more passive and indirect – in the background – in the case of small-group-based brand communities.

Firm-managed and customer-managed brand communities. In discussing the types of brand communities, another important distinction is between firm-managed and customer-managed brand communities. Many popular brands have many established brand communities, some organized and managed by the firm's professional marketing managers, and others that are grassroots organizations founded and run independently by customer enthusiasts. For example, Microsoft XBOX 360, the leading video gaming console, has an established brand community (www.xbox.com/Community) hosted on Microsoft's website and managed by Microsoft managers. Concurrently, there are dozens of XBOX brand communities, founded and managed by its fans such as the Brotherhood of the Box (www.bob.com.sg/forum) and Planetxbox360 (forums.planetxbox360.com).

There are several similarities between firm-managed and customer-managed brand communities. First, both communities are comprised of customers who are fans of the brand. Some participants may even have overlapping memberships within the two communities. Second, participants of both communities are interested in the same subject matter, namely news and information about the brand and its competitors. Finally, although consumers ultimately decide which community to join, there is at least some degree of targeting by community managers. For instance, it is quite common to offer referral rewards in cash, kind, or recognition, to existing members to recruit friends, and to use direct marketing approaches such as email invitations to encourage selected customer segments to participate.

Differences between firm-managed and customer-managed brand communities. There are also several significant differences between firm-managed and customer-managed brand communities. Perhaps the most important distinction has to do with the community manager's motives. Firms provide brand communities to their customers *to accomplish marketing objectives*. Some firms use brand communities to gather marketing research insights by monitoring discussions and/or interacting with participants. Others do so to increase participants' loyalty to their products and brands, and to increase their purchase behaviors. Consequently, customers are recruited to join the community through targeted approaches, and the facilities and affordances provided to participants are designed to reach these objectives. It is not uncommon for firms to use established segmentation variables, in particular, demographic variables for recruiting community participants. This professional recruitment of fans into communities often contradicts with the social motives of the brand's enthusiasts. This may be the reason why acquisitions and transformations of a customer-managed community into firm-managed communities often fail.

In contrast, customer-managed brand communities rarely have specific marketing goals. Instead, the community managers seek to express their love and admiration for the brand through organizing and managing the community. Participants, too, self-select and join the customer-managed brand community because of a shared passion for the brand, which often overlaps with common values, hobbies, and lifestyles. These differences indicate that customers in firm-managed brand communities should be more similar to each other in demographic characteristics, and those in customer-managed brand communities should be more likely to share psychographic commonalities.

The second difference between the two brand communities has to do with managers' constraints in the two cases. As noted earlier, professional managers of firm-managed brand communities are dictated by the firm's marketing objectives and therefore strive for consistency with its other marketing programs. Consequently, their emphasis is on a fit between the tone and content of the communications of the brand community members and the other marketing communications being sent by the firm. For example, a discussion in the brand community criticizing a newly introduced product and discussing its weaknesses will be viewed as inconsistent with the firm's ongoing advertising campaign that extols its virtues. Likewise, managers have an ingrained discomfort with giving customers free reign because of the possibility of unbridled and/or prolonged criticism by them regarding the firm. On the other hand, members of a customer-managed community do trust other members and their opinions more, because the firm's influence is not emergent. Practitioners and researchers are only now beginning to recognize and study these and other differences between firm-managed and customer-managed communities. This is one research area of great future potential, promising to increase our understanding of the scope and workings of brand communities.

Consequences of Consumer Participation in Brand Communities

Perhaps the most important reason for the success and growing importance of brand communities stems from the significant, multifaceted and long-lasting effects that participation in brand communities has on consumers. Consumer researchers have found that a key psychological process that occurs due to brand community participation is that the consumer psychologically identifies with the community, which in turn, mediates a number of important firm-relevant outcomes. It is useful, therefore, to understand what identification means, and examine its effects on the psychology and behavior of the brand community participants.

Identification with the brand community. Brand community identification captures the strength of the consumer's relationship with the brand community, whereby the person construes himself or herself to be a member – that is, as "belonging" to the brand community. In contrast to other personal identities, which may render a person unique and separate, this is a shared or collective identity. The consumer's self-esteem is also boosted to the extent that his or her ego-ideal overlaps with that of the others, and acting as the other acts or wants one to act reinforces one's self-esteem. Identification resembles aspects of normative and informational influence, as well as referent power, and is characterized by the community member's social identity. Several studies suggest that social identity, defined in terms of a valued group, such as a brand community, involves cognitive, affective, and evaluative components.

Considering the cognitive component first, identification with the brand community involves categorization processes, whereby the consumer formulates and maintains a self-awareness of his or her membership within the community (e.g., "I see myself as part of the community"), emphasizing his or her perceived similarities with other community members and dissimilarities with nonmembers. The self is perceptually and behaviorally depersonalized in terms of the relevant group prototype. The cognitive component of identification captures the consciousness-of-kind aspect of brand communities.

Next, the affective component of identification implies a sense of emotional involvement with the group, which social psychologists have characterized as an affective commitment to the group and which can also be viewed as kinship between members. Organizational researchers have shown that affective identification influences in-group favoritism and citizenship behaviors toward the organization. Identification means that the consumer agrees (or strives to agree) with the community's norms, traditions, rituals, and objectives, and promotes its well-being. Third, the *evaluative component of social identity – group-based self-esteem* – has been defined as the positive or negative value connotation attached to brand community membership, and arises from evaluations of self-worth derived from membership. Group-based self-esteem

has been found to promote actions that produce in-group welfare by social psychologists.

Identification with the brand community is a useful psychological concept to understand consumer psychology in this context because it produces behavioral intentions to engage in brand community participation, and to maintain a positive self-defining relationship with other community members. It also produces a number of interesting and important outcomes. Among the consequences of identification with the brand community are brand relationship quality, learning, oppositional loyalty, and trust in the firm.

Brand relationship quality. Consumer research on brand community has shown that the consumers' integration within the brand community is a function of their relationships with the brand, other community members, the product, and the firm as a whole. Brand relationship quality is the customer's psychological attachment to the brand and its assessment as a satisfactory partner in an ongoing relationship, and is consistent with the idea that consumers frequently view brands, especially well-liked or beloved ones, in human terms, assigning animate characteristics to them. Brand relationship quality captures the extent to which the consumer identifies with the brand and views his or her self-image as close to or overlapping with the brand's image. It involves cognitive aspects such as the degree to which the consumer believes that the brand's image overlaps with his or her self-image, and emotional elements such as the degree of the consumer's emotional attachment to the brand.

Brand relationship quality and brand community identification share a bidirectional relationship with one another. For some consumers, the consumer's relationship with the brand precedes and contributes to his or her relationship with the brand community. Many consumers first discover and value the brand for the functional and symbolic benefits it provides. A harmonious relationship with the brand can lead consumers to seek out and interact with like-minded consumers who share their enthusiasm. Moreover, an existing identification with the brand facilitates integration and identification with the brand community. For example, even when traditions, such as greeting other brand users, appear peculiar to the consumer, a strong relationship with the brand may help the person accept them and intrinsically endorse these practices.

On the other hand, integration with, and participation in the brand community, strengthens the customer's relationship with the brand. Not only does participation in the brand community provide opportunities for learning about the product and the brand associations (as we discuss below) but also close contact with other brand devotees rubs off on the consumer, increasing the strength of his or her emotional attachment to the brand. In today's environment, many marketers consider brand relationship quality to be the ultimate objective and metric of their marketing actions.

Learning. For complex, frequently evolving products, customers must continuously learn to keep abreast of changes and new developments and to take advantage of these advances. Traditionally, companies have been responsible for educating their customers to fully leverage the company's service offerings. They have done so by getting customers to participate in firm-organized training delivered by the firm's full-time or contracted employees. However, increasingly, many firms supplement this employee-based education model with peer-to-peer education delivered through brand communities, wherein customers assist their peers in learning about the products and their use. Customers thus take over service functions that are customarily performed by employees and act as "partial employees" of the firm. In eBay's Help Forums, for example, many novice sellers are interested in learning about the efficacy of different decision variables such as offering "buy-it-now" options, starting the auction with a low price versus a high price, and so on. Experienced sellers who are Forum members provide this service, bypassing the eBay employees.

The knowledge that a firm's customer possesses regarding its products, in particular, how to choose and use them, is a significant, valuable, and archivable resource, especially for complex products. Many brand communities exist to tap into, and they disseminate this

knowledge effectively to the participant base. For the individual member, this knowledge exchange translates into *learning*, which is defined as the customer's perceptions of increase in his or her own product expertise. Often, customers come to these communities for the first time because of the need to solve a specific problem. Once they receive the solution from the community, and hence have learned, they are captivated by the experience and become and stay a member themselves.

In brand communities, customers learn in vicarious and interactive ways, from the anecdotes, suggestions, and ideas of other members. The created knowledge is thus dependant on the members involved rather than on the pure truth. As members not only report short questions and/or manual-like answers but also share information on the situation, intentions, and feelings about an issue or action, they are able to convey the context. The context information contributes substantially to the learning experience, as it stimulates effective thinking, and supports information prioritization and interpretation. Additionally, in most brand communities, participants can access knowledge repositories such as records of prior conversations, product manuals, user guides, "hacks" from other members, and archives of (frequently asked questions (FAQs), that most communities store. The community site constitutes a knowledge base of all past incidents to aid the solutions of similar problems in the future and is therefore the collective memory of the individual interactions.

Oppositional brand loyalty. Prior consumer research has shown that brand communities strengthen a member's devotion and loyalty to the brand. For many loyal customers, the most important facet of this loyalty is derived from developing and expressing negative perceptions of competing brands. In fact, some researchers argue that this phenomenon of opposition to another brand and its community is the very defining feature of the brand community. *Oppositional brand loyalty*, which is defined here as "the participant's perception that competing brands are inferior to the target brand and should be avoided," benefits the firm by reducing the likelihood that members will purchase competing brand products, strengthening the possibility of future purchases of the firm's brand.

For the brand community participant, the tendency to position the brand against the competition arises from the perception of being threatened, along with a desire for the target brand to maintain its superiority over the competition. Such an adversarial position is fostered by the perceived normative pressure to conform to the brand community's views and to signal this inclination to conform explicitly to the world, and to dissociate from the disfavored brands through means visible to the community.

Trust in firm. *Trust in the firm* is defined as "the brand community member's willingness to rely on the firm, stemming from a confidence in its benevolence, reliability, and integrity." Trust in the firm is important from a managerial standpoint not only because it creates a relationship that is highly valued by the customer, extending beyond the interactions with the brand community, but also because it has been linked directly to organizational performance.

Research has shown that trust in the firm is affected positively by community identification because the shared consciousness inherent in greater social identification supports stronger convictions about the firm's intentions and integrity, leading to greater trust in it. This reasoning is consistent with the influential commitment-trust marketing theory, which posits that shared values between partners positively affect their trust perceptions. Trust in the community's manager and governance structure also contributes to the brand community member's trust in the firm. This is because the beliefs in the benevolence, reliability, and integrity of the brand community manager, along with the discussions in the trusted community regarding the firm's products and brands, elevate the participant's beliefs in the good intentions and reliability about the firm, leading to greater trust in it.

Conclusion

Brand communities offer the promise of a marketing program that is synergistic with the

intrinsic motivations, interests, and empowerment of contemporary consumers. Widely applicable and increasingly used by mainstream consumers across a range of product and service categories, they represent avenues for marketers to generate a range of positive outcomes for the firm in cost-effective ways. For consumer researchers, brand communities are venues to study a host of psychological and social issues. Brand communities are certain to grow in research importance and practical significance in the coming years.

Bibliography

Algesheimer, R. (2004) *Brand Communities*, Gabler, Wiesbaden.

Algesheimer, R. and Dholakia, U.M. (2006) Community marketing pays. *Harvard Business Review*, **November**, 26–28.

Algesheimer, R., Dholakia, U.M., and Herrmann, A. (2005) The social influence of brand community: evidence from European car clubs. *Journal of Marketing*, **69** (3), 19–34.

Algesheimer, R. and Gurau, C. (2008) Introducing structuration theory in communal consumption behavior research. *Qualitative Market Research*, **11** (2), 227–245.

Almeida, S.O., Dholakia, U.M., and Mazzon, J.A. (2009) The mixed effects of participant diversity and expressive freedom in firm-managed and customer-managed brand communities Working paper, Rice University.

Bagozzi, R.P. (2005) Socializing Marketing. *Marketing – Journal of Research and Management*, **1**, 101–111.

Bagozzi, R.P. and Dholakia, U.M. (2002) Intentional social action in virtual communities. *Journal of Interactive Marketing*, **16** (2), 2–21.

Bagozzi, R.P. and Dholakia, U.M. (2006a) Open source software user communities: a study of participation in Linux user groups. *Management Science*, **52** (7), 1099–1115.

Bagozzi, R.P. and Dholakia, U.M. (2006b) Antecedents and purchase consequences of customer participation in small group brand communities. *International Journal of Research in Marketing*, **23** (1), 45–61.

Cova, B. (1997) Community and Consumption: toward a definition of the 'Linking Value' of products or Services. *European Journal of Marketing*, **31** (3/4), 297–316.

Dholakia, U.M., Bagozzi, R.P., and Pearo, L.K. (2004) A social influence model of consumer participation in network- and small-group-based virtual communities. *International Journal of Research in Marketing*, **21** (3), 241–263.

Dholakia, U.M., Blazevic, V., Weirtz, C., and Algesheimer, R. (2009) Communal service delivery: how customers benefit from participation in firm-hosted virtual P3 communities. *Journal of Service Research*, **12** (2), 208–226.

Latour, B. (2005) *Reassembling the Social. An Introduction to Actor-Network-Theory*, Oxford University Press, Oxford.

McAlexander, J.H., Schouten, J.W. and Koenig, H.F. (2002) Building brand community. *Journal of Marketing*, **66** (1), 38–54.

Muniz, A.M. Jr. and O'Guinn, T.C. (2001) Brand community. *Journal of Consumer Research*, **27** (4), 412–432.

consumer neuroscience

Hilke Plassmann, Carolyn Yoon, Fred M. Feinberg, and Baba Shiv

Background

The past decade has seen tremendous progress in academic research at the nexus of neuroscience, psychology, business, and economics. Even five years prior to that, fewer than a half-dozen papers appeared with keywords "neuroscience" and "decision making." Presently, across these parent disciplines, the yearly count stands around 200, and is doubtless accelerating. The twin births of neuroeconomics and decision neuroscience has generated wide-ranging, ongoing debates on whether these hybrid fields benefit their parent disciplines and, within them, what forms these benefits might take (Shiv *et al.*, 2005). Their joint aim is to adapt tools and concepts from neuroscience – combined with theories, formal models, rich empirical data, and tested experimental designs from the decision sciences – to develop a neuropsychologically sound theory of how humans make decisions, one that can be applied to both the natural and social sciences.

A group of consumer psychologists is now dedicated to investigating consumer research questions with methodological and conceptual approaches from neuroscience. This emergent field, consumer neuroscience, is described in detail in this article. A primary, and critical, distinction is between "consumer neuroscience," which refers to academic research at the intersection of neuroscience, psychology and marketing, and "neuromarketing," which refers to practitioner and popular interest in neurophysiological tools – such as eye tracking, skin conductance, electroencephalography (EEG), and functional magnetic resonance imaging (fMRI) – to conduct company-specific market research. This article briefly details recent methods in neuroscience used by consumer researchers, presents basic ideas in consumer neuroscience as demonstrated by a variety of preliminary findings, and concludes with an outlook for the future of consumer neuroscience research.

Methodological Approaches in Neuroscience

Many distinct methods are used in neuroscience to study neural processes underlying human behavior. Because each method has its own strengths and weaknesses, robust research findings typically arise from studies using several different techniques to shed light on the same question. A core defining precept is that neurophysiological methods measure responses of either the central or the peripheral nervous system. Most research in the nascent field of consumer neuroscience has availed of methods capturing changes, or manipulating activity, in the central nervous system, specifically, in the brain. However, physiological measures are hardly new to consumer research; even 30 years ago, researchers had measured skin conductance and eye movements to understand motivation and involvement in consumer behavior. Many peripheral physiological reactions can be readily measured, and used to make inferences about both neural functioning and correlated behavior. For example, pupil dilation is correlated with mental effort; blood pressure, skin conductance, and heart rate are correlated with anxiety, sexual arousal, mental concentration, and other motivational states; and emotional states can be reliably measured by coding facial expressions and recording movements of facial muscles.

Recent technological advances in measuring and manipulating brain activity allow us to observe in real time the neural processes underlying consumer decision making via functional brain imaging techniques. Most brain imaging research involves within-subjects comparisons of people performing different tasks; an "experimental" task (A) and a "control" task (B). The difference between changes in brain activity measured during A and B indicates parts of the brain that are differentially activated by A (this is often referred to as the "subtraction approach"). One of the oldest imaging methods, EEG, measures electrical activity on the brain's surface using electrodes attached to the skull. EEG records timing of activity very precisely (resolution about 1 ms), but spatial resolution is poor, so that localizations when recording brain activity in subcortical areas that are small (e.g., the amygdala) can be problematic.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Positron emission topography (PET) is a newer technique that records positron emissions after a weakly radioactive blood injection. It does not measure brain activity directly, but rather metabolic changes linked to differentials in brain activity. PET offers better spatial resolution than EEG, but poorer temporal resolution and, because of rapid radioactive decay, it is limited to shorter tasks. However, PET usually requires averaging over fewer trials than fMRI, the method most widely used in consumer neuroscience. fMRI measures local changes in the ratio of oxygenated to deoxygenated hemoglobin. This ratio tracks neural activity because the brain effectively "overshoots" in providing oxygenated blood to active parts of the brain. Oxygenated blood has different magnetic properties than deoxygenated blood, giving rise to the signal picked up by fMRI (the so-called blood-oxygen-dependent-level, or BOLD, signal). Unfortunately, the signal-to-noise ratio of fMRI is, to date, fairly poor, so drawing tight inferences requires repeated sampling and many trials.

Yet another approach is single neuron recording, which tracks smaller scale neural activity (by contrast, fMRI measures activity of circuits consisting of thousands of neurons). In single neuron recordings, tiny electrodes are inserted into the brain, each measuring the firing of one specific neuron. Since these electrodes can damage neurons, the method is restricted to animal and special human populations, for therapeutic reasons (e.g., epileptic patients undergoing neurosurgery). Owing to its experimental use on animals, single neuron measurement has so far shed far more light on basic emotional and motivational processes than on higher-level ones, such as cognitive control. Regardless, the great body of extant research in neurobiology based on animal work (e.g., in rats and nonhuman primates) can directly inform theorizing in consumer neuroscience. Owing to functional and structural similarities in human and animal brains, the "animal model" has proved highly useful in the past, the main difference being a cortex enfolding the mammalian brain responsible for higher cognitive functions. Thus, owing to partial functional overlaps in subcortical areas, studying lower level processes, such as motivational signals during simple decision making, is also informative for understanding human decision making. An advantage of animal work is the ability to perform manipulations (e.g., stimulation) to make causal inferences, and also to allow single neuron recordings as direct measures of neuronal activity, which are not possible with fMRI or PET.

The oldest neuroscientific approach applied to understand human decision making, and among the cornerstones in decision neuroscience, is studying patient populations with brain lesions. Localized brain damage, often produced by accidents and strokes, and patients who underwent radical neurosurgical procedures, are an especially rich source of insights. If patients with known damage to area X perform a particular task more poorly than "normal" patients, this suggests that area X may be vital in performing that task. "Virtual lesions" can also be created by transcranial magnetic stimulation (TMS), which creates temporary local disruption to brain regions using magnetic field stimulations.

Concepts and Preliminary Findings

Consumer neuroscience evolved alongside wide-ranging developments in behavioral decision-making research and cognitive neuroscience, with the common goal to better understand various elements of consumers' evaluation and purchase decision processes (for a recent review, see Kenning and Plassmann, 2008). In consumer behavior research, neuroscience has received considerable attention for at least two reasons. First, neuroscience can be viewed as a new *methodological tool*, a "finer scalpel" to dissect decision-making processes without asking consumers directly for their thoughts, evaluations, or strategies. Second, neuroscience can be viewed as a source of *theory generation*, supplementing traditional ones from psychology and economics proper. Most of the remainder of this article is devoted to discussing these two perspectives.

Neuroscience as methodological tool. Methodological approaches in consumer research have tended to make heavy use of qualitative methods and survey measures to assess how experimental manipulations influence consumers' attitudes and behavior. This has served the field

well, having led to a rich body of empirical data and cohesive theoretical foundations. When relying on stimulus-organism-response models from psychology, consumer researchers must, however, take certain "black-box" conceptualizations of brain processes on faith. This raises several caveats that neuroscience might help to address.

First, neuroscience measurements, though they may be intrinsically noisy, have a strong advantage over surveys and self-reports in regard to potential biases. Since neuroscientific methods measure brain activity and its correlates directly – rather than relying on what subjects tell us what they think of how they are thinking it – they may offer more reliable indices of certain variables important to consumer researchers. Consider research on emotions and their role in consumer decision making. Emotions play an important role in consumer research, but are notorious for being difficult to induce via clever experimental manipulation, owing to their partially unconscious nature and great response heterogeneity across subjects. Neuroscientific research suggests that we cleave the concept of emotion into two parts, emotional states that can be measured through physiological changes (such as autonomic and endocrine responses), and feelings, the subjective and largely ineffable experience of emotions (Bechara and Damasio, 2005). The emotional states themselves depend on basic (implicit) brain mechanisms, which are rarely available for conscious cognitive introspection. A similar division could well be made for motivational processes. Another general area where neuroscience may offer substantial measurement benefits concerns when consumers undertake rapid information processing (e.g., viewing a visually dense TV ad) or enacting speedy habitual choices (e.g., selecting which sort of eggs or milk to buy at the supermarket). For example, prior research that used steady state, visually evoked potentials (similar to EEG) to understand memory systems underlying ad recall found that changes in brain activity in certain frontal brain regions, while viewing TV ads, predict long-term memory for those ads. Habitual choices were investigated by Milosavljevic *et al.* (2009), whose subjects engaged in a fast perceptual choice task between very familiar food items, during which eye positions were acquired with the help of an eye tracker. They found that subjects were able to make value-based choices (i.e., consistent with subjects' preferences) within "a blink of an eye" (as fast as 400 ms).

Second, some preliminary consumer research uses brain imaging to *validate marketing scales*. A guiding principle is that brain imaging is too costly a tool to squander on large-scale surveys, but could be used to ask the "right" questions to get at the underlying (neuro)psychological phenomena. For example, a recent study by Dietvorst *et al.* (2009) used fMRI to investigate neural activity in brain regions associated with "theory-of-mind" abilities (i.e., MPFC, temporo-parietal junction, temporal pole) in salespeople, combining it with surveys and other traditional methodologies to develop a new scale for assessing salespeople's interpersonal mentalizing skills.

Third, fortifying existing models of consumer decision-making with neuroscientific data may help them make *better predictions* about consumer behavior. An early effort in this direction is by Knutson *et al.* (2007), who combined neural and attitudinal measures to predict consumers' purchases. The authors decomposed the purchasing process into three steps – (i) viewing a product, (ii) viewing product and price information, (iii) pressing buttons to indicate whether one wishes to buy the product at the end of the experiment – and investigated neural correlates of the preference formation stages (i, ii) and the price processing stage (ii). They found that product preference correlated positively with activity changes in, amongst other areas, the nucleus accumbens (NAcc), a region thought to be involved in reward prediction mechanisms, and that net value (WTP-price) correlated positively with activity changes in the medial prefrontal cortex (MPFC), anterior cingulate cortex (ACC), and frontopolar cortex. During the choice stage (iii), purchasing correlated negatively with activation in the bilateral insula, a region known to be involved in risk and pain processing, and positively with activity changes in the ventromedial prefrontal cortex (VMPFC), a region shown to encode preference signals at the time of choice. When distinguishing purchased-item trials from non-purchased-item

trials, the authors found significant differences in NAcc activation during preference formation, and both MPFC and insula deactivation during price processing, in line with their *a priori* hypotheses. They then estimated brain activity in these three regions of interests and entered them as covariates in a logistic regression, along with self-report measures of preference and net value, to predict subsequent purchasing decisions. Results indicated that the full model (i.e., including the neural measures) was a significantly better predictor than one including only self-report measures.

This idea has been further developed by recent work combining behavioral decision-making research with machine-learning algorithms used in computational neuroscience. Computational neuroscience attempts to understand mental processing so as to allow a computer to mimic the way the brain functions during these processes. It has been used extensively to model simple learning algorithms in humans, among other areas. A recent first attempt to follow this path in consumer neuroscience (Tusche and Haynes, 2009) investigated how implicit brain processes could predict hypothetical purchasing decisions using multivariate decoding. They found that activity changes in the insula and the MPFC – specifically, while one group of subjects was exposed to various cars and asked how much they liked each (referred to by the authors as a condition where attention was shifted to cars, but not to purchasing) and while another group of subjects was asked to respond to a fixation cross that sometimes was displayed on a car background (referred to as condition with neither attention to cars nor to purchasing) – predicted at the end of the experiment whether or not subjects wished to purchase the cars in question.

A similar approach using eye-tracking data as physiological measures, coupled with models from computational neuroscience of vision, has started to be applied in advertising research; extensions relying on additional neurophysiological measures and more detailed computational models have recently been suggested by Milosavljevic and colleagues (e.g., Milosavljevic *et al.*, 2009). Indeed, studies of which brain areas (and/or other physiological measures, such as eye movements) are involved during certain tasks could well be enhanced by the use of machine-learning algorithms for predictive analysis.

Fourth, consumer researchers have begun to apply neuroscientific methods to test the abilities of *competing behavioral theories* to explain various phenomena. Although this has been the most common application of neuroscientific methods in consumer research to date, space limitations allow for only a few selected studies to be discussed. An early example is Yoon *et al.* (2006), who used fMRI to test whether semantic judgments about products and persons are processed similarly, finding that, contrary to several extant theories in marketing, they tend not to be. Specifically, when judgments of persons activated the MPFC, a region that in prior studies had been implicated in person processing, judgments of brands differentially activated the left inferior frontal cortex, an area known to be associated with object processing. Plassmann *et al.* (2008) used fMRI to study whether information that creates expectations about how good a product should taste (e.g., its price or brand) does so via postconsumption rationalizing or via changes in actual taste perceptions. The authors found the latter – that changing the prices of otherwise identical wines affected brain regions involved in interpreting taste pleasantness while the wines were being sampled. Hedgcock and Rao (2009) used fMRI to investigate different theories of how consumers make trade-offs between goods or services that differ on the utilities for single attribute, but whose overall utilities are similar, that is, are judged as equally good or bad across multiple dimensions. Prior research had confirmed that "asymmetric dominance" can lead to consistent violations of the regularity axiom: that introducing an alternative that is normatively irrelevant (because it is dominated by the existing alternatives) to the choice set can *increase* the choice probability of a nearby, dominating option. The authors found that activity patterns differed across conditions (e.g., higher in the dorsolateral part of the prefrontal cortex and the ACC vs higher in the amygdala, MPFC, and parietal lobule), supporting the existence of trade-off aversion.

Weber *et al.* (2009) investigated whether increased happiness, as opposed to alternative explanations consistent with rational choice

theory, can explain why consumers judge the value of money on the basis of actual amount of currency (nominal value) and not on the bundle of goods it can buy (real value), the so-called "money illusion." The authors found that brain areas thought to be involved in the anticipation and experience of reward, namely the ventral medial portions of the prefrontal cortex (PFC), showed higher activity changes when subjects displayed the money illusion. These findings were interpreted to suggest that the money illusion is based on changes in reward- or happiness-related neural activity, and thus cannot be fully accounted for by standard "homo economicus" theories of rational choice.

Two other recent papers investigated the neural basis of the "endowment effect": why we value goods we own more than (identical or equivalent) goods we do not. Vast theorizing has attended the endowment effect, ranging from a higher attraction to goods in one's possession (possibly owing to familiarity, or overestimating positive and underestimating negative features) to an aversion of losing what one tangibly possesses. Knutson *et al.* (2008) compared situations where subjects sold various products, bought different products, and made purchasing decisions for yet other products, all while their brains were scanned using fMRI. They found that, in both the selling and buying conditions, product preferences correlated with activity changes in the striatum (more precisely, the NAcc), a region known to be involved in reward prediction. The authors found no difference in NAcc activity in the selling versus buying conditions during what they refer to as the "product preference formation stage" (i.e., the point when subjects were exposed to items and prices), evidence against the theory that owned goods are more attractive or "sticky." In addition, the authors did not find activation in the insula (a region known to be involved in pain and risk processing, and so related to loss aversion) to correlate with product preference in selling versus buying trials. However, in "sell" trials specifically, individual differences in insula activity for preferred products *did* predict the extent to which subjects' indifference points for selling differed from the mean indifference point of buying (referred to by the authors as "endowment effect estimates"). The authors offer this as evidence of some role for insula activity, and thus for loss aversion as antecedent to the endowment effect. The second paper, De Martino *et al.* (2009), used fMRI to investigate the neural basis of within-subjects differences in WTA (willingness-to-accept)–WTP (willingness-to-pay) for lottery tickets, when subjects either owned (i.e., acted as seller) or did not own (i.e., acted as buyer of) the ticket. On a behavioral level, they found a systematic increase in the minimum selling prices, as compared to maximum buying prices, for a ticket with the same expected value. On a neural level, they found that the magnitude of WTP (in the buying condition) was encoded in the medial OFC (orbitofrontal cortex), and the magnitude of WTA (in the selling condition) in the lateral OFC. As the medial OFC has been found to encode increases and decreases in WTP during purchasing decisions across several fMRI studies (Plassmann *et al.*, 2007), and the lateral OFC was found to be responsive to the price (and thus the net value) of a good and to the anticipation of monetary losses, the authors suggest this finding as in line with the theory that, in the selling condition, transactions were perceived as potential losses. Interestingly, Knutson *et al.* (2008) found a partly overlapping area to be involved in the value comparison portion of the preference formation stage (i.e., only that time interval when subjects are exposed to the buying/selling price, a moment when subjects potentially started to think of their WTP/WTA in monetary units, rather than overall product preference, and compare it to the price, thus computing utility of the offer, i.e., whether or not it is a "good deal"). They found that the subjects showed increased activity changes in the VMPFC (among other areas) when they perceived the offer to be a good deal. However, it seems that the lateral part of the PFC did not react in the same way as in the study by De Martino *et al.* (2009), which could be owing to the fact De Martino *et al.* looked at neural correlates of within-subjects endowment effects, whereas Knutson *et al.* studied between-subject effects.

Of special note is that De Martino *et al.* (2009) also computed the difference between a "context-free" subjective value measure and a (either buying or selling) "context–biased"

one, and found the bilateral ventral striatum to correlate with increasing deviations from the unbiased value in the selling condition and decreasing deviations in the buying condition. The authors interpret this finding as evidence for the fact that the ventral striatum tracks the magnitude to which the subject's stated price deviated from the subject's true, unbiased value of the ticket in the selling and buying condition (i.e., reference dependent values) similar to a reward prediction error signal. Taken together, the findings of these two studies suggest that the endowment effect is related to negative emotional signals in the brain before and during the actual experience of endowment effects, supporting the loss aversion hypothesis consistent with prospect theory.

Neuroscience in Consumer Behavior Theory Generation. Consumer researchers have now begun to base hypotheses directly on *theories from neuroscience*. A recent example is Wadhwa *et al.* (2008), who investigated the impact of food sampling on subsequent consumer behavior, and compared two *a priori* hypotheses: the first – concordant with marketing practitioners, health experts, and much folk wisdom – that sampling a food will lead to lower subsequent consumption, and a second, rival hypothesis based on physiological theories of "reverse-alliesthesia," that as drive states affect the incentive value of relevant rewarding stimuli, a consumption cue high in incentive value (such as sampling a food) can strengthen drive states like hunger and, thereby, lead to an increase in the urge to engage in reward-seeking behaviors (such as eating more, or increasing other consumption-related behaviors). A series of behavioral experiments support predictions arising from the notion of reverse-alliesthesia – sampling a food or beverage items high in incentive value can in fact make individuals more likely to engage in reward-seeking behaviors, independently from specific reward type. Specifically, sampling a drink high in incentive value (e.g., Hawaiian Punch) not only leads to increased consumption of other drinks (e.g., Pepsi), but also to consumers giving higher desirability ratings for hedonic food, hedonic nonfood, and on-sale items, compared with those who had not sampled the high-incentive drink. Subsequent studies by the same authors investigate how motivational behavior impacts goal striving. In particular, they argue that if experiencing a hedonic cue enhances subsequent reward-seeking behaviors, then the induced motivational drive is also likely to enhance pursuit of a subsequent goal (defined as a representation of an internal state associated with a desirable outcome). For example, experiencing a hedonic cue (e.g., being exposed to romantic pictures) is likely to make one persist longer on a subsequently adopted intellectual goal of solving anagrams. Thus, unlike much of the extant research that has focused on how factors related to the goal state (e.g., desirability) can influence its pursuit, the focus here is on how factors *un*related to the goal state (e.g., incidental brief experiences with hedonic cues) can enhance subsequent goal pursuit. The authors' hypotheses are based on neuroscientific evidence (in rats and humans) of how the dopamine system works, that is, any hedonic cue that leads to enhanced dopamine activity could also motivate behaviors aimed in pursuit of a subsequently adopted goal associated with a desirable outcome.

Can cognitive neuroscience benefit from consumer research? Having detailed why consumer researchers are increasingly interested in joining forces with neuroscientists, let us briefly suggest how they might repay the debt, that is, "what's in it for neuroscientists?" In neuroscience, a number of developments led to a "cognitive revolution" that set the stage for the field of *cognitive* neuroscience, specifically. One major concern in early neuroscientific work was that it was largely descriptive in nature, and led to multiple isolated theories resistant to integration into a general, normative theory. An important advance came via the introduction of signal detection theory, a first attempt to relate neuronal activity directly to behavior in the field of vision science. We believe it is possible that sophisticated *formal models* from decision science and economics might help cognitive neuroscience establish a body of normative theory regarding how different types of decisions are enacted in the brain. A similar approach is currently used in so-called "model-based fMRI" studies, which investigate

neural correlates of decision-making variables by integrating models from economics and behavioral decision to aid in statistical data modeling.

Several pioneering neuroscientific studies on emotion and decision making were conducted using lesion patients, discovering impairments during simple economic decision-making tasks for patients with lesions specifically in the ventromedial portion of the PFC (for a review, see Bechara *et al.*, 2000). This represented something of a milestone for interdisciplinary work at the intersection of neuroscience, psychology and behavioral decision science, and stimulated a great deal of subsequent research. An opportunity for neuroscientists in this area is the possibility to combine the *wealth of empirical data* and *formal experimental design methods* from decades of work in behavioral decision science and studies on lesion patients. Although tasks have to be adapted to the specific requirements of patient work, experimental and empirical findings about observed behavior can be readily transferred to patient studies. This potential of behavioral decision research for neuroscience becomes slightly more complicated for brain imaging studies using fMRI, owing to the requirement for repeated measures. Nevertheless, behavioral protocols and empirical knowledge from behavioral decision science have greatly benefited the study of neural correlates of decision making. A prime example is a study of Plassmann *et al.* (2007) that used a design from behavioral economics, the Becker-de-Groot-Marchack auction, to sample trial-by-trial incentive compatible and nonhypothetical psychometric measures (here, economic preferences in the form of WTP bids) to correlate in real-time with repeated neurometric measures (here, changes in BOLD-signal). The authors found that activity in the medial orbitofrontal cortex and dorsolateral prefrontal cortex correlate with the magnitude of the subjects' WTP.

Summary and Future Directions

In this article we introduced the genesis, core concepts, and preliminary empirical findings of the nascent field of consumer neuroscience. The future of the field, and its eventual reception, will depend on the insights and benefits it can generate in concert with its parent disciplines. We believe it will be crucial that researchers within the field of consumer neuroscience adopt a multimethod approach, including not only different neuroscientific tools but also traditional behavioral (laboratory) and field experiments, to transcend the limitations of mere correlational results subject to inverse inference and causation. Ideally, consumer neuroscience research will be able to link hypotheses about specific brain mechanisms (location, activation, direction, connectivity) to both unobservable intermediate variables (utilities, beliefs, goals, etc.) and observable behavior (such as choices), using a variety of different methodological approaches from neuroscience, statistical modeling, and social science proper.

Two emergent trends deserve special mention. First, computational, model-based consumer neuroscience studies will become increasingly crucial, as statistical approaches themselves become ever more sophisticated, from mapping functional connectivity to the use of multivariate statistics to decoding algorithms from computational neuroscience for predictive analysis. Second, a recent development of clear relevance to consumer neuroscience is the study of individual differences based on genetic information (often referred to as "imaging genetics"). Imaging genetics is the study of how genetic differences lead to individual differences in the morphology and functions of the brain, and thereby differences in behavior. The rapid proliferation of inexpensive, personalized genetic information should make such studies increasingly accessible and informative for consumer researchers.

The old saw goes that prediction is difficult, especially about the future. And so it is with consumer neuroscience, whose large time, financial, and learning-curve costs may appear to place it beyond the current reach of many consumer behavior researchers. We anticipate that the inevitable waning of these barriers over the coming decade will produce a flowering of interest in understanding the neural bases and correlates of consumer behavior, and encourage our fellow researchers to wade enthusiastically into this exciting, growing area of inquiry.

See also *consumer decision making; consumer information processing; implicit consumer cognition*

Bibliography

Bechara, A. and Damasio, A.R. (2005) The somatic marker hypothesis: a neural theory of economic decision making. *Games and Economic Behavior*, **52**, 336–372.

Bechara, A., Tranel, D. and Damasio, H. (2000) Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain*, **123** (11), 2189–2202.

De Martino, B., Kumaran, D., Holt, B., and Dolan, R.J. (2009) The neurobiology of reference-dependent value computation. *Journal of Neuroscience*, **29**, 3833–3842.

Dietvorst, R.C., Verbeke, W.J.M.I., Bagozzi, R.P. *et al.* (2009) A Sales Force– Specific Theory-of-Mind scale: tests of its validity by classical methods and functional magnetic resonance imaging. *Journal of Marketing Research*, **46** (5), 653–668.

Hedgcock, W. and Rao, A.R. (2009) Trade-off aversion as an explanation for the attraction effect: a functional magnetic resonance imaging study. *Journal of Marketing Research*, **46**, 1–13.

Kenning, P.H. and Plassmann, H. (2008) How neuroscience can inform consumer research. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **16**, 532–538.

Knutson, B., Rick, S., Wimmer, G.E. *et al.* (2007) Neural predictors of purchases. *Neuron*, **53**, 147–156.

Knutson, B., Wimmer, G.E., Rick, S. *et al.* (2008) Neural antecedents of the endowment effect. *Neuron*, **58**, 814–822.

Milosavljevic, M., Huth, A., Rangel, A. and Koch, C. (2009) Ultra-rapid Consumer Choices. California Institute of Technology working paper.

Plassmann, H., O'Doherty, J. and Rangel, A. (2007) Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *Journal of Neuroscience*, **27**, 9984–9988.

Plassmann, H., O'Doherty, J., Shiv, B. and Rangel, A. (2008) Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 1050–1054.

Shiv, B., Bechara, A., Levin, I. *et al.* (2005) Decision neuroscience. *Marketing Letters*, **16** (3–4), 375–386.

Tusche, A. and Haynes, J.D. (2009) Brain signals reveal implicit consumer choices. Neuroimage, Special Issue Proceedings of the Organization for Human Brain Mapping Conference.

Wadhwa, M., Shiv, B. and Nowlis, S.M. (2008) A bite to whet the reward appetite: the influence of sampling on reward-seeking behaviors. *Journal of Marketing Research*, **45**, 403–413.

Weber, B., Rangel, A., Wibral, M. and Falk, A. (2009) The medial prefrontal cortex exhibits money illusion. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 5025–5028.

Yoon, C., Gutchess, A.H., Feinberg, F. and Polk, T.A. (2006) A functional magnetic resonance imaging study of neural dissociations between brand and person judgments. *Journal of Consumer Research*, **33**, 31–40.

self-regulation

Kathleen D. Vohs and Ayelet Fishbach

People are not passive, listless automatons waiting for stimuli to demand a response. They plan, organize, prioritize, and modify their behavior. They set goals. Broadly, these behaviors are called *self-regulation*. Self-regulation is a vital process of the self. The self-regulation process not only sets goals but also monitors progress, shifts to new goals as necessary, and manages a hierarchically structured system of multiple goals.

The consumption setting is an ideal context in which to situate self-regulation theory and research. After humans evolved to form groups that led to the formation of culture, the development of trade and the emergence of humans as *homo economicus* brought about advances in reproduction and survival, not to mention quality of life (*see* CONSUMER WELL-BEING).

The role of consumption is still crucial today but not in the same way. The press to study self-regulation in contemporary settings comes in large part from the abundance of temptations and urges. "Having it now" interferes with long-term goal strivings, many of which are centered around consumption. Financial goals influence spending and health goals influence food intake (Herman and Polivy, 1975). It is difficult to understate the importance of self-regulation in humans ancestral history and now.

Definitions

Terms: Cognitive scientists use the construct of basic level to describe the default category used to define objects. In the current case, the basic level constructs are self-regulation and self-control. Many scholars use these terms interchangeably to mean modulation of an incipient response. Often the definition tacitly includes a secondary component that a more suitable response will be substituted for the undesirable response. Modulations of responses can involve suppression (the most common) or amplification. It can even mean creating a response out of nothing (e.g., faking a smile when opening a disappointing gift). When scholars use the terms differently, they often take self-control to be the conscious form of response modification, whereas self-regulation can be nonconscious as well. At a broader level, some scholars have discussed self-management, which involves moving current goals up and down the priority list, the importance or salience of the goals often being determined situationally.

Scope: There are five general classes of responses that can be regulated. One can fixate one's attention (e.g., when driving in a Minnesota snowstorm), down- or up-regulate emotions (e.g., stifling a laugh), overcome an impulse (e.g., to eat the whole bag of cookies), suppress intrusive thoughts (e.g., do not think about an ex-lover), or guide behavior (e.g., throwing a dart to hit the bull's-eye). Eating, drinking, spending, sexuality, drug taking, avoiding exercise, and risky driving involve many or all of these foundational spheres of self-regulation and are but a few of the manifestations of self-regulation in a consumer culture.

Types: The current article cleaves the literature on self-regulation into two pieces: motivational versus social-cognitive inputs to self-regulation. It is no coincidence that one major distinction between types of self-regulation, namely controlled versus automatic processes, roughly corresponds to the degree to which they are energy-taxing (i.e., subject to motivation) and conscious. The automatic system guides behavior via well-learned (through practice or habit) or functional (i.e., evolutionary-based) action patterns. Owing to their ingrained nature, these responses are not said to take up energy or to be under conscious control. In contrast, consciously controlled processes cost energy to enact and involve a high degree of self-awareness. Hence, essential distinctions about broad types of self-regulation processes rest on the notion of motivation and cognition.

Process versus outcome: We emphasize the fact that self-regulation is a process, not an outcome. Too often, scholars conflate the process of attempting to reach a goal with the normative nature of the goal itself. Mostly people set goals that align with normative prescriptions for behavior, such as having

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

positive relationships with others, adopting healthy behaviors, and being morally good, but not always. People can engage in self-regulation in order to get themselves to overeat, binge drink, or engage in risky sex acts. More novel, insightful, and predictive research will come from separating the process of self-regulation from the outcomes that the actors intend to achieve.

Overview

As mentioned, the article reviews research on self-regulation as falling into one of two camps: energy or motivational versus social-cognitive. Whether a theory focuses on self-regulation from a motivational or cognitive perspective suggests that different inputs, processes, and outcomes are considered relevant in each approach. Admittedly, though, the dividing line between these two camps is artificial, if readers think that one theory or another could be in a different camp they may be right. We do not intend to draw exact dividing lines. Rather, parsing the research into motivational versus cognitive emphases is meant to be a loose classification system in order to highlight theories that share features.

Motivational Inputs to Self-Regulation

Motivation can loosely be defined as effort or energy that one puts toward a goal (*see* MOTIVATION AND GOALS). Freud's psychodynamic model has been the energy model with the greatest impact on the psychology of self-change. As the reader is probably noting already, that model has long since dominated scientific studies. Baumeister, a social psychologist who studies self-regulation, once quipped that, "energy models are so out of fashion that we [i.e., social scientists] aren't even against them anymore" (personal communication, March 16, 2004). Today's motivation research, while certainly not Freudian in nature, embraces energy as a central component that reorients behavior away from entrained routes and toward desired ends.

Cybernetic models. Cybernetic models describe the process of self-regulation as discrepancy-closing. When a person identifies a desired end state, the motivational system calculates the size of the discrepancy between the present state and the desired state and guides action toward closing the gap. The acronym "TOTE", which stands for Test, Operate, Test, Exit, is often used to denote this process. According to this notion, once the person identifies a desired end state, the required effort to reach this state is assessed (Test), which leads people to put effort into achieving it (Operate), which requires another assessment of the distance (Test), which cycles around recursively until the process ends because the end state is achieved (Exit). For example, a woman may perceive she needs to go on a diet. She calculates how much weight to lose, trims calories and exercises, and steps on the scale from time to time. She stops dieting when the gap is closed either because she reached her goal (success) or altered it (likely due to repeated failed attempts).

Research by Carver and Scheier (1990) developed this feedback model into a comprehensive model of feedback loops. Their model highlights emotions as feedback for self-regulation. Positive emotions signal that the rate of closing the gap to goal attainment is faster than expected. In this case, people reduce their effort or "coast." Negative emotions, conversely, signal that the rate of closing the gap is slower than intended. In this case, people increase their effort investment. An interesting implication of this model is that people will work harder toward a goal when they feel bad about it than when they feel good about it.

Limited-resource model of self-regulation. A recent model depicts self-regulation as being a function of a limited stock of energy. This energy is said to be involved in every act of self-regulation, which implies that it is easily taxed. When people are low in self-regulatory resources, they are said to be in a state of depletion (or ego depletion), which portends poor self-regulation subsequently. At last count, over 120 published experiments have shown the pattern that was predicted by the limited-resource model: after a person engages in self-regulation, ensuing attempts at self-regulation are less successful than if the person had not earlier engaged in self-regulation. The resource is renewable but unfortunately it

does not seem to replenish itself with as much ease as it gets depleted.

The first papers on the limited-resource model established basic self-regulation findings (Baumeister *et al.*, 1998). In this work, participants would engage in one form of self-regulation (or not, for the neutral conditions) and then all participants would be tested in a different domain of self-regulation. Papers subsequent to those focused on boundary conditions of the model or mechanisms. Particularly germane to the current volume is work on breaking one's diet and impulsive spending as a function of self-regulatory resource depletion and decision making impacting self-regulation (*see* IMPULSIVE AND COMPULSIVE BUYING).

Controlling caloric intake is a perennial self-regulation problem and one that is especially interesting given that people cannot completely exit the eating cycle. That is, when studying other overconsumption problems, such as drug or nicotine addiction, one solution is to simply not consume the substance. This does not work with food intake. Hence, how people grapple with regulating eating is of special interest because people must indeed constantly regulate.

The limited-resource model was used to show that dieters, but not non-dieters, eat more when they have been taxed of their self-regulatory resources (Vohs and Heatherton, 2000). In a representative experiment, for instance, dieters and non-dieters were brought to the laboratory individually and were first asked to watch a boring movie about bighorn sheep. There were also snacks in the room that were either nearby the participant's chair or far away, and those snacks were either labeled as "off-limits" or "for the taking." Being bored made it likely, we thought, that participants would want to eat the snacks, but we knew from prior research that dieters may well be tempted but would not eat the snacks because they were fattening (chocolate candies, cookies, etc). In fact, we hypothesized that the greatest temptation would be for the participants for whom the snacks were nearby and allowed to be eaten. In this condition, not eating the goodies would require the most self-regulation. For dieters, we surmised that if they were reduced in the amount of self-regulatory resources they had, then later if we surprised them with another context in which they had to taste food (a taste-and-rate ice cream task), they would not be able to exert enough resources to curb their desire to indulge. We observed the expected pattern. As predicted, dieters who sat next to candies and were allowed to eat them – but held back because of their personal dieting goal – later ate the most ice cream during a ratings task. Non-dieters are not regulating in the domain of food intake, so their eating – in the initial phase and the ice cream eating test phase – is merely a function of internal cues and not because of depletion. This article raises a broad point about the limited-resource model of self-regulation and motivation: when people become depleted, they will not behave impulsively at random. What will come unglued is that which they are regulating the most: for dieters, this means that they should eat more when depleted, which are the findings of Vohs and Heatherton (2000). For consumers trying to curb their spending, this would mean purchasing. Vohs and Faber (2007) found exactly this pattern.

The studies testing impulsive spending as a result of self-regulatory resource depletion (*see* IMPULSIVE AND COMPULSIVE BUYING) are similar to the aforementioned studies on eating among dieters with one key difference. We argued in this article (Vohs and Faber, 2007) that everyone has an outside goal of reigning in spending (*see* CONSUMER MATERIALISM). In almost all situations, some self-control in buying is necessary in order to achieve other goals. People who think they have unlimited funds, such as celebrities, may soon discover that their funds have run dry (e.g., Mike Tyson's bankruptcy, Michael Jackson's debts). Hence, we claimed that most people have at least an implicit goal not to spend impulsively. This generalization meant that we expected to see unplanned spending (as a form of lax self-regulation) as a consequence of self-regulatory resource depletion as a main effect. Above and beyond this main effect, we also predicted that the buying behavior of people who are regulating in the domain of spending would be especially affected by resource depletion. As predicted, our experiments revealed that depletion condition and interaction of depletion condition with trait impulsive buying tendencies

are major causes of impromptu spending. One representative experiment had participants in the ego depletion condition list out their thoughts for several minutes, with the exception that they were not allowed to think thoughts of a white bear. Other participants were allowed to think anything they wanted for the thought listing task, including a white bear. Hence, the former group has a thought suppression goal, whereas the latter does not. After the thought listing task, participants were shown 22 products that were said to be under consideration for inclusion in the university bookstore. We gave them $10 to spend in the mock bookstore or take home. The dependent measure was ad hoc spending in this spontaneous purchasing situation. As expected, participants who had been in the thought suppression condition spent more money and bought more items than participants in the no suppression condition. Moreover, participants who had chronic problems with overspending spent more than other participants when they had been depleted of their self-regulation resources. In short, controlling one's spending uses self-regulation, and when self-regulatory resources run low, unplanned and unnecessary spending is likely to occur.

The third domain that is especially relevant to the notion of consumer self-regulation is how choice influences self-regulation. The executive function of the self houses both self-regulation and decision making. As a consequence, making choices and controlling the self are likely to be interrelated (*see* CONSUMER DECISION MAKING). We posited that making choices would tax the self-regulation system in a manner similar to when people engage in self-regulation. That is, we predicted that self-regulatory abilities would be compromised after making choices. Eight studies supported this hypothesis (Vohs *et al.*, 2008).

In one study, participants in the choice condition (comprised primarily of first-year college students) were asked to select courses that they would take for their remaining years in college, whereas those in the no-choice condition simply reviewed the course catalog. Next, all participants were given time to study for an upcoming test purportedly to measure participants' intelligence. Participants in the choice condition practiced less (that is, procrastinated more) than those in the no-choice condition. In another experiment, participants in the choice condition were asked to make a series of binary choices between consumer products, such as a yellow candle versus white candle. Participants in the no-choice condition rated the same products on how often they had used such products but did not make choices. Then all participants were asked to keep their forearms in painfully cold water for as long as possible. As predicted, high choice participants were less able to endure the pain of cold water than no-choice participants. In a study that took place at a shopping mall, shoppers reported the extent they had made decisions during the day, including how many decisions, how much they deliberated and how personally responsible they felt for the decisions, which formed a latent variable that we called investment in choices. Later, shoppers were asked to complete arithmetic problems. The more choices shoppers reported having made, the fewer math problems they attempted, the less time they spent on them, and the more problems they got wrong.

Follow-up studies found that when people are depleted of their self-regulatory resources, they are more likely to buy hedonic products more than utilitarian products. These studies also confirmed the linkages between self-regulation and choice in a non-North American context (Western Europe), which is important because one would want to know that these effects occur outside of the United States and its plethora of choices. In sum, these studies suggest that the processes involved in decision making rest on the same mechanism that is used for self-regulation, which unfortunately exact significant intrapsychic costs.

Other Motivational Inputs

Motivation has much to do with whether the actor thinks that he or she is capable of reaching the goal. One consumer model of goal attainment (Bagozzi and Dholakia, 1999) focuses on self-efficacy, which is the belief that a person can achieve what he or she wants to achieve. Self-efficacy is a construct somewhere between self-esteem (one's appraisal of oneself in terms of goodness or badness) and self-regulation (goal attainment) in that it involves self-perceptions

of ability. Having the sense that one can meet valued goals is crucial to self-regulation success because it amps up the motivation to do so.

A small but growing literature on lay beliefs about the structure of self-regulation also points to the role of motivation. At a broad level, personal theories of self-regulation as a flexible, dynamic system has been found to benefit self-regulation.

Social Cognitive Inputs to Self-Regulation

Social cognitive models vary by whether they address the regulation of a single versus multiple goals. In addition, models that address regulation of multiple goals vary by whether they address goals that are of equal status (e.g., goals related to career and family) versus goals that pose a self-control dilemma (e.g., eating healthy and tasty). We offer a brief overview of some of the more prominent models in this domain.

Regulatory focus theory. Other research distinguishes between the different types of goals that individuals pursue and how they may evoke different processes of self-regulation. The most basic distinction exists between the goal to obtain pleasure versus avoid pain. For example, every organism strives to get food as well as avoid danger (and stay alive). On the basis of this distinction, regulatory focus theory distinguishes self-regulation with a promotion focus, on the presence and absence of positive outcomes or gains, and self-regulation with a prevention focus, on the presence or absence of negative outcomes or losses. Individuals strive toward positive outcomes when pursuing advancement needs and aspirations (e.g., planning a vacation). They further strive to avoid negative outcomes when pursuing security needs and responsibilities (e.g., getting a fire alarm). Interestingly for many goals (e.g., personal hygiene), both frames are possible and individual differences as well as situational factors will determine which focus a person adopts.

According to regulatory focus theory, different processes characterize the different foci and they have different psychological consequences. In particular, the pursuit of promotion goals is oriented toward receiving gains, whereas the pursuit of prevention goals is oriented toward non-losses. The emotional consequences of pursuing promotion goals further differ from those of prevention goals. For promotion goals, successful pursuit results in happiness and failure results in sadness, because these emotions characterize the presence versus absence of gains. In contrast, for prevention goals, successful pursuit results in calmness and failure results in anxiety, because these emotions characterize the absence versus presence of losses. In addition, research finds that people enjoy pursuing a goal in a manner that sustains their regulatory focus, a phenomenon titled "regulatory fit" (Higgins, 2000). In particular, people prefer to use eager strategies toward promotion goals and vigilant strategies toward prevention goals.

Goal systems theory. Whereas the basic unit of research on self-regulation involves the pursuit of a single goal, people rarely hold only one goal at a time. Rather, they often need to juggle between several goals that compete with each other for resources (e.g., buying a house versus increase contribution to retirement) or that directly undermine each other (e.g., eating healthy and fatty food). Research on goal systems theory addresses the cognitive operations that govern the regulation of multiple goals. According to this research, goals are organized in associative networks, connecting higher order goals with lower level means of attainment. Each goal is associatively linked to several attainment means. Similarly, each means is connected to several goals that it can potentially serve. For example, the goal of having a good figure may be connected with the means of eating healthy food and exercising, and the means of exercising may be connected with goals of having a good figure and leading a healthy lifestyle. In addition, there are inhibitory links between unrelated goals or between unrelated means that potentially compete for resources.

There are several predictions that follow from this theory. For example, it assumes that the presence of potential means to a goal activates the corresponding goal, such that individuals will adhere to the goal upon encountering a means (or an opportunity) for goal pursuit. For example, the mindless impact of an advertisement on

consumer choice may reflect such goal activation upon encountering cues for means of attainment. In addition, because competing goals inhibit each other, a focal goal can often lead to the inhibition of another alternative goal. In this way, the focal goal "*shields*" itself from alternative ones by directly reducing their accessibility (Shah, Friedman, and Kruglanski, 2002). Empirically, this inhibition is often reflected in the slowing down of lexical decision times to concepts that represent alternative goals. For example, activation of academic goals slowed down the lexical decision time to concepts related to alternative goals (e.g., exercising).

Other research on goal systems documents a desire for multiple goal attainment. Given the presence of several salient goals and limited motivational resources, individuals search for attainment means that are linked to the attainment of several goals simultaneously. For example, a person may prefer to dine out (vs. dine in) in order to satisfy both hunger and various social motives (to see and be seen, etc.), or commuters may choose to commute by bike (vs. car) in order to save money and keep in shape. Interestingly, however, as the number of goals attached to a given means increases, each association becomes weaker. The result is a dilution of the means–goal association, which may reduce the perceived effectiveness of the means with respect to the goal. That is, although individuals often seek means that are linked to multiple goals, they also find them less effective than those which serve fewer goals. In a demonstration of this *dilution effect*, Zhang, Fishbach, and Kruglanski (2007) found that when participants considered the different goals (e.g., building muscles and losing weight) that a single means (e.g., working out) could satisfy, an increase in the number of goals resulted in a reduction in the perception of the instrumentality of the means with respect to each goal. As a result, individuals sometimes choose means that are connected with fewer goals, because they believe these means are more effective. For example, participants in a study were less likely to use the writing function of a pen that had also been used as a laser pointer (vs. was not used as a laser pointer) when they only needed to write. It appears that the general preference for means that serve several goals diminishes and even reverses when a person focuses upon pursuit of a single goal.

Dynamics of Self-Regulation

When individuals wish to successfully pursue several goals, they often take into account a sequence of actions that unfold over time. Then, completed actions in the past or upcoming actions in the future can both influence which goal an individual attends to in the present. In addition, missing actions in the past or the absence of plans to pursue a goal in the future will influence the decision to pursue a goal in the present.

Research on the dynamics of self-regulation (Fishbach and Dhar, 2005) attests that self-regulation follows a sequence of actions that either balances between several goals or highlights the pursuit of one, focal goal. According to this research, whether people balance or highlight depends on whether they represent actions that pursue a goal as an expression of their commitment to this goal or as a signal that progress was made. When actions express commitment, each action increases goal commitment and goal adherence, and hence people highlight the goal. However, when actions signals progress on a goal, each action decreases the perceived need for taking more actions, and hence people balance between their goal pursuits.

One factor that determines whether people represent goal actions as an expression of their commitment or as making progress is their degree of commitment certainty. When individuals are not sure whether the goal is valuable for them, they highlight by adhering more to a goal if they have done so before. But when individuals are already sure that the goal is valuable, they wish to monitor their level of progress and they adhere to a goal more if they have not done so since they perceive greater need to make progress. For example, shoppers who were buying luxuries, expressed greater interest to use a frequent buyer card that had a visual emphasis on completed purchases (using stamps) than on missing purchases (using punches), because completed purchases signaled that buying luxuries is valuable and promoted highlighting of the purchase behavior. In contrast, shoppers who were buying necessities were more likely

to use a frequent buyer card if it emphasized missing (vs. completed) purchases, because missing purchases signal need for progress and promote balancing for the absence of purchases.

These findings have implications for pursuing group goals, when individuals join force with other group members, for example, when contributing to a charity. Research finds that first time donors are more likely to contribute if they receive information on accumulated contributions to date (vs. missing contributions to go), whereas returning donors are more likely to contribute if they receive information on missing contributions to go (vs. accumulated to date). The reason is that first time donors wish to evaluate whether the charity goal is valuable and their behaviors follows, or "highlights" others' contributions. Returning donors wish to evaluate the rate of progress on the charity goal and their behavior compensates or "balances" for others' contributions.

The arrangement of action alternatives also influences the dynamic of self-regulation individuals follow. Individuals often make selections from sets that include items that serve multiple goals. For example, they browse a television guide that includes educational shows and light sitcoms, or they go through highbrow news magazines and lowbrow fashion magazines on a newsstand. The arrangements of the alternatives influences people's perceptions of them as competing against versus complementing each other, which in turn influences the dynamics of self-regulation they would follow. In particular, separating items into two sets (e.g., two bowls), versus presenting them together in one set (e.g., one bowl), determines whether individuals perceive the items as conflicting versus complementary. When the items are presented apart, they seem conflicting and promote a highlighting dynamic of choice; when the items are together, they seem complementary and promote a balancing dynamic of choice.

These dynamics have unique consequences for situations in which the items on a set pose a self-control conflict. When goal and temptation alternatives (e.g., healthy and unhealthy foods) are presented apart from each other, they seem to compete against each other. As a result, people are more likely to resolve the conflict in favor of the goal alternatives in a dynamic of highlighting: they assign a greater value to goal alternatives (e.g., educational shows, news magazines) than to tempting alternatives and consistently choose goal alternatives for both immediate and future consumption. In contrast, when choice alternatives appear together and seem to complement each other, thus promoting balancing, people tend to resolve the self-control conflict in favor of the immediately gratifying temptation option. As a result, they value the tempting alternatives (e.g., watching sitcoms, reading lowbrow fashion magazines) more than the goal alternatives and prefer these tempting alternatives for immediate consumption, thereby postponing the consumption of goal alternatives to a future occasion. The reason tempting alternatives are selected first in this presentation format is that their value is immediate, whereas the value of the goal alternatives, although larger, is delayed. Thus, in a self-control conflict, a balancing dynamic would most often take the form of "first temptation then goal" rather than "first goal then temptation."

Counteractive control theory. Other cognitive models address the exercise of self-control and overcoming temptations. In particular, counteractive control theory (Fishbach and Trope, 2007) examines the processes of resolving goal conflict between an important long-term goal and a momentary temptation (e.g., saving and spending, eating healthy and enjoying fatty food, study and leisure). According to this theory, the process of counteracting temptations involves asymmetric shifts in motivational strength, namely an increase in the motivation to pursue a goal and a reduction in the motivation to pursue temptations. Such asymmetric shifts are often of conscious, deliberative nature, but they may involve nonconscious, implicit strategies that promote individuals' long-term interest without requiring conscious awareness. In addition, some strategies are behavioral and they act on the choice opportunities themselves (e.g., increasing the availability of goal items) while other strategies are more cognitive by nature, and they act on the representation of the choice opportunities (e.g., increasing the value of goal items).

Individuals employ behavioral counteractive strategies when they choose, for example, to

make rewards contingent on undergoing uncomfortable (but helpful) medical tests and make penalties contingent on failing to do so. Using this strategy, they counteract the temptation to avoid the test. Another behavioral strategy involves precommitment to pursue goals and avoid temptations. For example, while grocery shopping, a consumer might suspect that having tempting sweets available in the kitchen will pose a problem and decide to purchase fruits instead. People further avoid temptations by maintaining physical distance from tempting objects, while ensuring proximity to objects associated with goals. For example, they will push away the liquor glass and pull toward them a glass of water, in order to stop drinking.

Other cognitive strategies alter the representation of goals and temptations. Thus, individuals often counteractively bolster their evaluations of goals and dampen their evaluations of temptation. For example, health-conscious individuals who were facing a choice between health bars and unhealthy chocolates evaluated the chocolates as less appealing than the health bars before choosing between the two, in order to secure making the "right choice." Individuals further promote goal pursuit by adopting a concrete representation of goals but an abstract representation of temptations, because the more concrete representation facilitates action. For example, in a study on the regulation of academic goals, students formed concrete behavioral plans to facilitate pursuit of their academic goals (Gollwitzer and Brandstätter, 1997). In another study on delay of gratification, children resisted the temptation to eat a marshmallow by thinking about it as an abstract cloud, thus cooling its appetitive influence (Mischel, Shoda, and Rodriguez, 1989).

Finally, similar patterns of asymmetric shifts exist for individuals' nonconscious self-control operations. For example, they nonconsciously boost the value of the goal while dampening the value of the temptation. In addition, implicit counteractive control entails changes in the accessibility of goals and temptations. Individuals shore up their goals by activating goal-related constructs in response to interfering temptations and by inhibiting temptation-related constructs in response to goal-related cues. For example, a study on college students demonstrated that subliminally presenting the word "television" reduced the time students took to subsequently recognize the goal-related word "study" (goal-related), and similarly presenting the word "study" increased the time students took to subsequently recognize the word "television" (Fishbach, Friedman, and Kruglanski, 2003).

Conclusions

The research reviewed in this article is aimed at offering two broad perspectives on self-regulation. One is motivational, the other cognitive. For the former, energy models are increasingly popular. Future work could, however, expand on the link between emotion and motivation as these two often interact in novel ways (e.g., urges and impulses as emotions; *see* CONSUMER DESIRE).

We encourage researchers to remember that self-regulation is a process, not an outcome. While it is true that most self-regulation acts are aimed at securing goals that will lead to normatively better, longer, and morally right lives, some do not. An emphasis on the process rather than the outcome will open new vantage points.

Problems because of failures of self-regulation seem to be taking over the Western (and, increasingly, Eastern) world. Drug and alcohol abuse, overeating, sexual improprieties, underachievement, insufficient exercise, Internet and video game addictions, violence, and crime are all linked through problems with self-regulation. Indeed, it is difficult to think of many behavioral problems that do not have a regulatory component. Against this backdrop, it is somewhat consoling that along with problems of self-regulation comes an increasing emphasis on the importance of research in this area.

Bibliography

Bagozzi, R.P. and Dholakia, U. (1999) Goal-setting and goal-striving in consumer behavior. *Journal of Marketing*, **63**, 19–32.

Baumeister, R.F., Bratslavsky, E., Muraven, M., and Tice, D.M. (1998) Ego depletion: is the active self a limited resource?. *Journal of Personality and Social Psychology*, **74**, 1252–1265.

Carver, C.S. and Scheier, M.F. (1990) Origins and functions of positive and negative affect. *A Control-process View. Psychological Review*, **97**, 19–35.

Fishbach, A. and Dhar, R. (2005) Goals as excuses or guides: the liberating effect of perceived goal progress on choice. *Journal of Consumer Research*, **32**, 370–377.

Fishbach, A., Friedman, R.S., and Kruglanski, A.W. (2003) Leading us not unto temptation: momentary allurements elicit overriding goal activation. *Journal of Personality and Social Psychology*, **84**, 296–309.

Fishbach, A. and Trope, Y. (2007) Implicit and explicit mechanisms of counteractive self-control, in *Handbook of Motivation Science* (eds J. Shah and W. Gardner), Guilford, New York, pp. 281–294.

Gollwitzer, P.M. and Brandstätter, V. (1997) Implementation intentions and effective goal pursuit. *Journal of Personality and Social Psychology*, **73**, 186–199.

Herman, C.P. and Polivy, J. (1975) Anxiety, restraint and eating behavior. *Journal of Abnormal Psychology*, **84**, 666–672.

Higgins, E.T. (2000) Making a good decision: value from fit. *American Psychologist*, **55**, 1217–1230.

Mischel, W., Shoda, Y., and Rodriguez, M.L. (1989) Delay of gratification in children. *Science*, **244**, 933–938.

Shah, J.Y., Friedman, R., and Kruglanski, A.W. (2002) Forgetting all else: on the antecedents and consequences of goal shielding. *Journal of Personality and Social Psychology*, **83**, 1261–1280.

Vohs, K.D., Baumeister, R.F., Schmeichel, B.J. *et al*. (2008) Making choices impairs subsequent self-control: a limited resource account of decision making, self-regulation, and active initiative. *Journal of Personality and Social Psychology*, **94**, 883–898.

Vohs, K.D. and Faber, R.J. (2007) Spent resources: self-regulatory resource availability affects impulse buying. *Journal of Consumer Research*, **33**, 537–547.

Vohs, K.D. and Heatherton, T.F. (2000) Self-regulatory failure: a resource-depletion approach. *Psychological Science*, **11**, 249–254.

Zhang, Y., Fishbach, A., and Kruglanski, A.W. (2007) The dilution model: how additional goals undermine the perceived instrumentality of a shared path. *Journal of Personality and Social Psychology*, **92**, 389–401.

consumer information processing

Hélène Deval and Frank R. Kardes

Marketing researchers and marketing managers spend enormous resources on monitoring, predicting, understanding, and influencing the behavior of consumers. This requires investigating the effects of marketing communications at each stage of consumer information processing: attention, comprehension, evaluation, memory, and choice (Wyer, 2008). Consumers can use virtually any type of cue, attribute, or knowledge as information, and this information is processed sequentially to facilitate judgment and decision making.

Attention

Consumers are limited information processors (Kardes *et al*., 2011). They often lack the motivation, capacity, or opportunity to analyze the judgmental implications of all relevant pieces of information. It is therefore critically important to capture their attention and interest. Specific goals influence the amount of processing and lead consumers to focus selectively on certain products or services. A given product or service is valued more heavily when it is goal relevant (valuation effect) and consumers tend to ignore goal-irrelevant products (devaluation effect, Markman and Brendl, 2005). The devaluation effect is typically much larger than the valuation effect. For example, even money is devaluated when hunger, thirst, or some other need is highly salient, even if money could indirectly help satiate that need.

Regardless of their relevance to a specific goal, some stimuli are conspicuous and therefore capture attention automatically (Kardes, 1994). Specifically, salient and vivid stimuli are highly noticeable and difficult to ignore. The vividness of a stimulus depends on the properties of the stimulus itself. It is determined by concreteness, emotional interest, and psychological proximity, independent of the context or background (e.g., the color red is a vivid color). Salience is determined by a difference in the background – contextually novel, unexpected, complex, or moving stimuli stick out (e.g., a red package surrounded by gray packages on a shelf is salient but the same package becomes nonsalient if surrounded by other red packages). Not surprisingly, advertisers are well aware of the importance of capturing the attention of consumers and so constantly try to create novel and salient ads that capture consumers' attention and interest. For example, ads that have a surprise element, such as mystery ads that hide the identity of the product till the end of the ad, are attention drawing (Kardes *et al*., 2011). For unfamiliar brands, mystery ads help build a strong category-brand association in memory because withholding the identity of a brand arouses curiosity. This strong category-brand association results in the activation of a brand whenever the product category is brought to mind. Similarly, ads that use multiple variations on a theme, and ads that use upward camera angles are contextually salient and capture interest. In the same way, unusual packages, such as cylinders for Pringle's potato chips, plastic eggs for L'eggs panty hose, and 15-packs for Strohs beer, are salient and attention drawing. Salience effects are reduced when involvement or motivation for accuracy is high (Kardes *et al*., 2011).

Comprehension

Comprehension is the ability to interpret and assign meaning to information by relating it to knowledge already stored in memory (Wyer, 2008). It involves the formation of inferences that enable consumers to grasp the meaning of a product claim without systematically considering and evaluating each possible meaning of each word presented in the message (Kardes *et al*., 2008). For example, instrumental inferences (the statement "he pounded a nail into the wall" implies that a hammer was used) facilitate communication. Interpretive inferences usually involve very little effort, making it difficult for the consumer to discriminate between the information that was actually presented and interpretative inferences based on this information (Kardes, 1994).

Prior knowledge leads to selective processing of knowledge-consistent information. Take two people that have opposite opinions about computers – one is a PC user and the other

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

is a Mac user. It is likely that the PC will perform better on some dimensions (e.g., it is cheaper) and the Mac will perform better on other dimensions (e.g., it is more reliable). Both computer users will use prior knowledge and will selectively process the information that supports their original opinion, and, consequently, the same message will be interpreted in different ways by different consumers.

Similarly, most consumers assume that price and quality are highly correlated, and this assumption leads consumers to focus on information that is consistent with this hypothesis (Kardes *et al.*, 2008; Posavac *et al.*, 2005). Consequently, recognition and recall performance is better for high-price/high-quality brands and low-price/low-quality brands than for brands that do not confirm the original hypothesis. This leads to overestimation of the strength of the relation between price and quality. This overestimation increases the willingness of consumers to spend. Overestimation is reduced only when selective processing is discouraged owing to low cognitive load, random information presentation, or a low need for cognitive closure. The need for cognitive closure refers to a preference for a definite answer to a judgmental problem, any answer rather than ambiguity, inconsistency, or confusion (Kruglanski and Webster, 1996). Selective processing, anchoring, priming, primacy, heuristic reasoning, and stereotyping effects decrease as the need for cognitive closure decreases.

When processing effort is low, consumers are more prone to the effect of misleading advertisements using claims that are literally true but figuratively false. The claim "Brand X may help whiten your teeth" is literally true: the product may or may not help whiten your teeth. Yet, because in everyday language *may* is interpreted as *usually*, the claim is figuratively false. Likewise, consumers are prone to infer that the claim "Women who look younger use Oil of Olay" indicates that "Women who use Oil of Olay look younger" (Kardes, 1994). This is known as "confusion of the inverse," which occurs frequently in everyday reasoning.

Low processing effort also results in other types of biases, like the truth effect (Kardes, 1994). The perceived validity of a statement increases with mere repetition independently of its objective validity. Because consumers might be unwilling or unable to judge the validity of a statement directly, they use other cues such as familiarity to assess the validity of a product claim. If people feel that they have heard the claim before, they assume that the claim is probably true even if they cannot remember where they heard it in the first place. The truth effect leverages this relation between familiarity and perceived validity: repetition increases familiarity, and in turn, familiarity increases perceived validity. The truth effect has obvious implications with respect to the effects of repetitive advertising on the perceived validity of product claims. This effect is less pronounced when processing is high because consumers are more likely to recognize that the impression of familiarity is because of repetition.

Evaluation

After a product has captured consumer attention and interest, and after the information about the product has been understood, consumers assign a value to it in order to determine how much a product is worth to them (*see* ATTITUDES; PERSUASION). Selective processing leads consumers to focus on positive aspects of positively evaluated products and negative aspects of negatively evaluated products (Kardes *et al.*, 2011). Returning to our example of the PC user and the Mac user, selective processing has implications for the evaluation of products. The Mac user evaluates Macintosh computers positively and PCs negatively. He will therefore focus on the positive aspects of a Mac. Conversely, he will focus on negative aspects of a PC.

Selective processing can also induce product attribute information distortion. In this case, the information that does not support an initial preference is not ignored but its value is simply discounted while the importance of information supporting the initial preference is enhanced. When more than one product is present and consumers are encouraged to identify the product that is the leader, one alternative becomes the focal alternative. When consumers begin to prefer tentatively one alternative over others, the attributes of the

focal alternative are rated as more important or as more favorable, relative to attribute ratings assessed in a pretest or a control group. One piece of information that would otherwise be interpreted as neutral might be interpreted as supportive of the leader. Information distortion can amplify the pioneering brand advantage or increase preference for an inferior alternative when attribute information is presented in a manner that leads to an early preference advantage for that alternative.

Most of the biases described previously leverage the lack of involvement and/or motivation to carefully process information (*see* CONSUMER INVOLVEMENT). Fortunately, consumers do not always process information using heuristics and low effort. Important decisions are likely to be guided by judgments and evaluations formed on the basis of a systematic analysis. Dual process models, such as the elaboration likelihood model (ELM) or the heuristic systematic model, suggest that there are two qualitatively different routes to persuasion: a high involvement route in which consumers think a lot (i.e., the central route of the elaboration likelihood model and the systematic route of the heuristic/systematic model) and a low involvement route in which consumers think very little (i.e., the peripheral route of the elaboration likelihood model and the heuristic route of the heuristic/systematic model; Chaiken and Trope, 1999).

The ELM suggests that when involvement is high, and when the ability to think about an ad is high, consumers are likely to follow the central route to persuasion by focusing on information that is most central to or important for forming an accurate attitude. Consumers are less prone to process the information passively. Instead, consumers are likely to elaborate on a persuasive message by generating their own support arguments or counterarguments. Strong arguments and strong reasons for forming a particular attitude are particularly persuasive when consumers follow the central route to persuasion. Attitudinal judgments formed on the basis of highly compelling arguments are likely to be enduring, resistant to change, and more likely to influence behavior.

When involvement is low or when the ability to think about an ad is low because of distraction, a lack of relevant knowledge, time pressure, and so on, consumers are likely to follow the peripheral route to persuasion without evaluating the quality of the arguments supporting or refuting a given position. Following the peripheral route to persuasion by focusing on peripheral or superficial information makes it easy to form an opinion without much thought. Many different types of superficial cues may be used. Attractive, likable, and expert sources seem trustworthy, so if these sources say that an advertised brand is a good product, consumers often believe these sources rather than spending a lot of time thinking about the attributes and benefits of the product (*see* SOCIAL INFLUENCE). Furthermore, positive moods and feelings often transfer to the advertised brand when consumers follow the peripheral route to persuasion. Consumers tend to form weak attitudes that are not accessible from memory, not long lasting, not resistant to change, and that have a weak influence on other judgments and behavior.

The ELM indicates that facts and reason are important when consumers follow the central route to persuasion, but not when consumers follow the peripheral route to persuasion. The model also indicates that celebrities, authority figures, humor, and pleasant background music and scenery are important when consumers follow the peripheral route to persuasion, but not when consumers follow the central route to persuasion. Hence, consumers use different types of information depending on which route to persuasion they are following. Furthermore, the central route to persuasion leads consumers to form strong attitudes that are accessible from memory, persistent, resistant to change, and that have a strong influence on other judgments and behavior (Haugtvedt and Kasmer, 2008; *see* CONSUMER INTENTIONS).

MEMORY

Availability refers to information that is stored in memory, and accessibility refers to the "activation potential" of the available information (Wyer, 2008). Accessibility depends on the strength of the association between stored knowledge and situational cues, the recency with which information has been acquired or last activated, the frequency of prior activation, and

the intensity with which information has been processed (*see* KNOWLEDGE ACCESSIBILITY). Rather than using all relevant information stored in memory, consumers frequently use the subset that has been activated or primed by stimulus cues. The priming paradigm consists of two phases: first, consumers are exposed to a stimulus that increases the accessibility of the primed construct; second, consumers perform a seemingly unrelated task and their judgments or behaviors are assessed. When the applicability of a priming stimulus relative to a target is high, the stimulus influences subsequent judgments and behaviors concerning the target.

A wide variety of marketing communications have been shown to serve as priming stimuli (e.g., ads, Internet messages, salesperson interactions, retail environments, consumer magazines), and these priming stimuli have been shown to influence a wide variety of judgments (e.g., expensiveness, evaluative, and likelihood judgments) and mind-sets (e.g., comparing brands in one product category can increase purchase intentions in a different product category). Consumers often overestimate the likelihood of occupations (e.g., doctors, lawyers), objects (e.g., luxury cars, swimming pools), and behaviors (e.g., having wine with dinner, incidence of crime) shown frequently on television. This is particularly the case for heavy viewers, despite the fact that most consumers do not believe that television reflects reality accurately. With respect to evaluation, priming can result in either assimilation (a shift in judgment of a target toward the priming stimulus) or contrast (a shift in judgment of a target away from the priming stimulus), depending on the degree of overlap between the priming stimulus and the target. Assimilation occurs when overlap is high, and contrast occurs when overlap is low.

There is usually a considerable gap between the point of time at which information about products or services is encountered and the point of time at which this information is used in order to make a purchase decision. That is why it is of crucial importance to understand the role of memory in decision making (*see* CONSUMER MEMORY PROCESSES). Memory enables past experiences and learning that occurred long ago to influence current behavior. Memory is often compared to a computer (Kardes *et al.*, 2011). A computer has a hard drive that can store a large amount of inactive files. A computer also allows users to retrieve a file from the hard drive and make it active so that the file can be processed (e.g., edited or used). Similarly, people have a long-term memory system that stores a large amount of inactive data or knowledge. To use this knowledge, people must retrieve it from long-term memory and bring it in into short-term, active memory so that the information can be processed. The transfer from the short-term memory to the long-term memory happens through rehearsing the information. The entire thinking and reasoning occurs in short-term memory, and only a small amount of information can be held in short-term memory; consistent with Miller's magic number seven, consumers cannot hold more than seven brands in their short-term memory (Kardes *et al.*, 2011). If this information is not used, it can be lost in as little as 18 seconds. By contrast, long-term memory appears to store an unlimited amount of information permanently.

Memory can have an impact on diverse dimensions of brand choice: the brands included in the consideration set, what information will be used to evaluate the considered brands, and how the information will be used. Three general types of brand choice can be distinguished (Lynch and Srull, 1982). First, stimulus-based choice involves the selection of one brand from a set of brands that are actually present at the time of choice. For example, consumers in a grocery store might occasionally consider the brands that are on the shelves directly in front of them, not considering any brand retrieved from memory. Purely memory-based choice involves electing a brand from a set of brands stored in memory. Good examples of this type of choice include picking a restaurant before one leaves home. In between these two alternatives, mixed brand choice involves both alternatives that are physically present and alternatives that are retrieved from memory. This type of choice happens when consumers compare products from different stores that carry different brands (Kardes, 1994).

The usual comparison of memory to a computer does not imply that memory is infallible. For example, forgetting occurs. Forgetting is not because of information loss

or decay. Instead, forgetting occurs because we cannot find the information we are trying to find during memory search. In addition, information held in long-term memory can also be distorted or changed over time (Kardes *et al.*, 2011).

Selective processing also influences how information is stored in memory. Advertising can change the manner in which products are experienced. Subjective product experiences are ambiguous and might be interpreted in multiple ways. When an ad states that a particular brand of soda tastes sweet and refreshing, consumers frequently test this hypothesis by searching for hypothesis-consistent information in their memory (i.e., confirmation that the soda is indeed sweet and refreshing), and this leads to premature confirmation of the advertisement claim. When the ad precedes the experience, it sets expectations that will guide the interpretation of the product experience. Learning from experience becomes difficult because prior beliefs and current experiences are perceived as more consistent than they actually are. Conversely, current belief might bias memory. For example, a consumer might be led to believe that product A is better than product B. But, if at a later time, the consumer is led to believe product B is better than product A, the consumer is likely to distort his memory and report that he thought that product B was superior all along. Selective search can occur for stimulus information or for information stored in memory, depending on the timing of the ad and the consumption experience.

Most retail settings offer only a limited set of brands, and the brands that are offered often determine which brands are included in consumers' consideration sets, or the group of brands that are considered for purchase by consumers. Consumers are more likely to focus on immediately available brands and neglect unmentioned brands in singular judgment task (Kardes *et al.*, 2011).

The neglect of unmentioned brands increases as the number of brands presented increases because of part-list cuing-induced inhibition (Kardes *et al.*, 2011). This type of inhibition is different from intentionally blocking or suppressing examples available from memory. It occurs when just some brands are presented and consumers are trying to recall as many brands as possible. The brands presented inhibit the recall of other brands. Similarly, advertising reduces the ability of consumers to recall attribute information pertaining to competing brands. Inhibition of nonfocal brands also occurs when focal brands are physically present (as in grocery stores) rather than mentioned only by name. When no brands are mentioned by name, consumers are forced to generate their own consideration sets via memory search, and this leads to a strong advantage to brands enjoying strong brand-category associations in memory. Consumers often demonstrate selective memory for previously chosen products.

Choice

All the previous stages of information processing (attention, comprehension, and memory) are important in order to make a purchase decision (i.e., a choice). Selective processing often leads consumers to focus on one brand. Leveraging the effect of part-list cuing, marketers often try to influence which brands and how many brands consumers will include in their consideration sets. Marketers also leverage other effects, such as the attraction effect and the compromise effect (Kardes *et al.*, 2011; *see* CONSUMER DECISION MAKING).

The attraction effect is based on the principal that a given brand seems more attractive when it is compared to inferior brands and less attractive when compared to superior brands. Hence, marketers want consumers to compare their brands to inferior brands. This can be done in a selling situation or using comparative advertising and promotion where the target brand is shown to be superior to other brands on some dimension. Alternatively, the attraction effect can come into play when a firm has multiple brands in its product line. For example, when a company markets a relatively expensive product, it might appear as too expensive when presented with no point of comparison. As a result, few consumers might choose to purchase the product. If the company introduces a more expensive product that competes directly with the first product, the original product might appear as more reasonably priced and much more attractive. This gives the impression to consumers that they save

money while the characteristics of the original product remain unchanged.

Another way to improve the evaluation of a target brand is to make this brand seem like an average, or a good compromise, relative to other brands in the consideration set. A compromise brand seems average on all important attributes or features; other brands often have some really good features and some really bad features, but a compromise brand does not seem to have any very bad features and is at least acceptable on all features. Hence, the compromise brand seems like a safe choice and the compromise effect, or the increased likelihood of buying a compromise brand, is particularly likely to occur when consumers are concerned about avoiding a bad decision. Brands that are intermediate or "average" in terms of price, quality, and number of features are frequently chosen from the consideration set, and they are chosen even more frequently when we have to justify our choice to others. Because it is often easier to justify the purchase of intermediate rather than extreme brands, such a choice might also reduce post-decision regret. Of course, almost every brand can appear to be a compromise brand depending on what brands serve as points of comparison. Advertising and promotion campaigns that encourage consumers to compare a seemingly average brand to more extreme brands increase the influence of the compromise effect.

Selective processing leads consumers to use a number heuristics to make choices (Kardes *et al.*, 2011). An important aspect of making a purchase decision is based on the ability to predict the future performance of a product. Although the ability to make predictions is very important, research evidence suggests that people are not very good at making them. This is because people often unknowingly use heuristics to generate predictions. Such heuristics include the representativeness heuristic, the availability heuristic, the simulation heuristic, and anchoring-and-adjustment heuristic.

People using the representativeness heuristic make predictions based on the perceived similarity between a specific target and a general category. Consumers can make mistakes when they focus on superficial similarities between an object and a category (e.g., focusing on irrelevant similarities like package color and graphics). People using the availability heuristic make predictions based on the ease with which instances are retrieved from memory. For example, if an event is highly memorable and easy to remember (a service failure), it might overpower the availability of other less memorable events (e.g., several flawless interactions with the same service provider). If it is easy to remember, it seems more likely to occur. This will affect the prediction of consumers if they have to interact again with the same service provider. People using the simulation heuristic make predictions based on the ease with which an event or a sequence of events can be imagined or visualized. Such a heuristic has important implications for health-care marketing. Consumers with high blood pressure often forget to take their medicine because the symptoms usually are not that bad and this makes it difficult to imagine that their condition is serious. People using the anchoring-and-adjustment heuristic make predictions based on a first impression or an initial judgment (or anchor) and then shift (or adjust, or fine-tune) this judgment upward or downward depending on the implications of the imagined possibilities. Unfortunately, we usually fail to adjust enough and the effect of the anchor can remain more impactful than it ought to be. As a consequence the final judgment is often close to the initial impression.

In addition to heuristic that influence predictions related to decision making, there are also heuristic strategies related specifically to consumer choice (Kardes *et al.*, 2011). These heuristics include the lexicographic heuristic, the elimination-by-aspects heuristic, the additive-difference heuristic, the conjunctive and disjunctive heuristics, and the frequency of good and bad features heuristic. Consumers using the lexicographic heuristic (or single-attribute heuristic) compare all brands on one attribute, and choosing the brand that performs the best on that single attribute, generally ignoring the other attributes. If there is a tie, consumers examine the next most important attribute to break the tie.

Another choice heuristic is the elimination-by-aspects heuristic. Consumers simply reject brands that do not have a key feature that they want. Next, consumers would focus on a different attribute and reject all brands that do not meet their requirements on this attribute, and so on. Consumers using the additive-difference heuristic compare two brands at a time (rather than all brands) on one attribute at a time and subtract the difference. Subtraction is performed on all relevant attributes and each attribute is weighted for importance – each difference is multiplied by the importance of the attribute. Each weighted score is then summed to arrive at an overall score for each brand.

Choice heuristics do not always involve making comparisons across brands. Sometimes we focus on one brand at a time and no comparisons are performed. Examining one brand at a time is easy, and if the first brand seems satisfactory you might buy it without examining any other brands. Consumers using the conjunctive heuristic set a minimum value for all relevant attributes and select the first brand that meets this value for all relevant attributes. Consumers using the disjunctive heuristic set an acceptable value, rather than a minimum value, for all relevant attributes and select the first brand that meets this value on one attribute, rather than all of the attributes. The frequency of good and bad features heuristic is a strategy in which consumers form a simple attitude toward each brand alternative by counting the number of good product features, counting the number of bad product features, and choosing the brand with the largest number of good product features.

Conclusion

Consumer judgment and decision making is influenced by a large and complex set of variables that capture the attention and interest of consumers; influence how consumers acquire, retain, and revise product knowledge; and influence how product knowledge is used to ultimately make judgments and decisions. Selective processing is likely to have an impact on every stage of consumer information processing (e.g., selective attention, selective encoding, selective retrieval). Dual process models emphasize that effortful information processing, which reduces selective processing, heuristic processing, and the biases resulting from selective or heuristic processing, depends on the motivation and ability to process information carefully. Consumer information processing has important implications for understanding how consumers make judgments and decisions, and how policymakers can help consumers make better judgments and decisions.

Bibliography

Chaiken, S. and Trope, Y. (eds) (1999) *Dual-Process Theories in Social Psychology*, Guilford, New York.

Haugtvedt, C.P., Herr P.M. and Kardes F.R. (eds) (2008) *Handbook of Consumer Psychology*, LEA/Psychology Press, New York.

Haugtvedt, C.P. and Kasmer, J.A. (2008) Attitude change and persuasion, in *Handbook of Consumer Psychology* (eds C.P. Haugtvedt, P.M. Herr and F.R. Kardes), LEA/Psychology Press, New York, pp. 419–459.

Kardes, F.R. (1994) Consumer judgment and decision processes, in *Handbook of Social Cognition*, Vol. 2 (eds R.S. Wyer and T.K. Srull), Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 399–466.

Kardes, F.R., Cronley M.L. and Cline T.W. (2011) *Consumer Behavior*, South-Western College Publishing, Cincinnati.

Kardes, F.R., Herr P.M. and Nantel Jacques (eds) (2005) *Applying Social Cognition to Consumer-Focused Strategy*, Lawrence Erlbaum Associates, Mahwah.

Kardes, F.R., Posavac, S.S., Cronley, M.L. and Herr, P.M. (2008) Consumer inference, in *Handbook of Consumer Psychology* (eds C.P. Haugtvedt, P.M. Herr and F.R. Kardes), LEA/Psychology Press, New York, pp. 165–191.

Kruglanski, A.W. and Webster, D.M. (1996) Motivated closing of the mind: 'Seizing' and 'freezing'. *Psychological Review*, **103**, 263–283.

Lynch, J.G. Jr., and Srull T.K. (1982) Memory and attentional factors in consumer choice: concepts and research methods. *Journal of Consumer Research*, **9**, 18–37.

Markman, A.B. and Bendl, C.M. (2005) Goals, policies, preferences, and actions, in *Applying Social Cognition to Consumer-Focused Strategy* (eds F.R. Kardes, P.M. Herr and J. Nantel), Lawrence Erlbaum Associates, Mahwah, pp. 37–51.

Posavac, S.S., Fitzsimons, G.J., Kardes, F.R. and Sanbonmatsu, D.M. (2005) Implications of selective processing for marketing managers, in *Applying Social Cognition to Consumer-Focused Strategy* (eds F.R. Kardes, P.M. Herr and J. Nantel.), Lawrence Erlbaum Associates, Mahwah, pp. 37–51.

Wyer, R.S. (2008) The role of knowledge accessibility in cognition and behavior: implications for consumer information processing, in *Handbook of consumer psychology* (eds C. Haugtvedt, P.M. Herr and F.R. Kardes), Psychology Press, New York, pp. 31–76.

the role of schemas in consumer behavior research

Karen E. Flaherty and John C. Mowen

Introduction

Social and cognitive psychology researchers assume that because our information processing capabilities are limited, we develop abstract knowledge structures. One form of knowledge structure that has been extensively studied in both the psychological and consumer behavior literatures is the schema. A schema is defined as a "*. . . cognitive structure that represents knowledge about a concept or type of stimulus, including its attributes and the relations among those attributes*" (Fiske and Taylor, 1991, p. 98). Schemas are an individual's preconceived notions or ideas about the world that are stored in memory.

The schema concept stems from a combination of Heider's (1958) balance theory and Asch's (1946) person perception research, wherein people are believed to form general holistic perceptions from discrete social encounters (Fiske and Taylor, 1991). The basic premise of schema-driven research is that in order to simplify reality, we develop categories (*see* CONSUMER CATEGORIZATION) in which to store information at a broader level rather than attempting to incorporate specific experiences and new information in a piecemeal fashion. In essence, schemas serve to ease our information management processes and social experiences. Categories are formed to help us discriminate among individuals, interpret information, and evaluate others. When evoked, schemata are believed to influence social perceptions and evaluations as well as behavior. While schema research focuses largely on understanding and describing the result of the evoked schema on such outcomes, categorization research focuses on how instances are classified and categories (or schemas) are formed. In this article, we limit our discussion to schema research.

Types of Schemas

Various types of schemas have been studied across a number of disciplines, including social psychology, consumer behavior, marketing, and organizational behavior. For instance, categories and schemas have been proposed to capture individual's perceptions of themselves, other people, roles and occupations, and the situations they face. A person (or other) schema represents preconceptions of a variety of personality traits, social goals, and so on. For example, an individual may have a preconceived idea of what it means to be outgoing or competitive (e.g., Cantor and Mischel, 1979). A self-schema represents perceptions of the self, or the core of the self-concept (Markus and Wurf, 1987). Self-schemas tend to be more complicated than other schemas, and as a result people remember better self-relevant information (Kihlstrom *et al*., 1988). Role schemas capture our understanding of the characteristics and behaviors that we expect of people in certain roles and positions, including both achieved (e.g., professor, graduate student, janitor) and ascribed (e.g., race, age, gender). Event schemas represent our expectations regarding how events ought to be sequenced – much like scripts (e.g., what should happen when you go to a wedding, a funeral, or a dinner party). These types of schemas and others that have been proposed in the literature all serve to guide our information processing and social encounters with greater efficiency and ease.

The Effects of Schemas

In general, people balance a schema stored in memory against actual evidence received in the environment. Thus, external information is weighed against internal preconceived notions. In addition, schemas affect how individuals encode, remember, and evaluate information that is encountered after an initial schema is evoked. When a new piece of information or a new encounter takes place, it evokes an existing schema. In turn, the schema impacts the person's attention to the information, how the person interprets the information, and the basic judgments that are formed from the information.

Attention. Schemas affect what people notice. When an individual has a strong preconceived notion of a given subject or object (i.e., the schema is well developed), then the person is likely to pay close attention to schema-consistent information and to ignore schema-inconsistent

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

information (Fiske and Neuberg, 1990). Schemas also minimize perceptions of differences across those within the schema (e.g., all scientists are alike) and increase perceptions of differences between schemas (e.g., scientists and janitors are very different).

Memory. Schemas also affect what people remember (*see* CONSUMER MEMORY PROCESSES). In addition to its effects on attention, research suggests that when an individual has a well-developed schema, he or she is also more likely to remember schema-consistent information (Fiske and Neuberg, 1990). However, if an individual does not have a well-developed schema, then both consistent and inconsistent information will be processed and remembered. The individual is still attempting to form a solid perception of the category and as a result will attempt to integrate the new information.

Evaluations. Immediate evaluative responses are a result of the affect triggered by categorization to a specific schema (Anderson and Cole, 1990). When actual data or evidence is encountered, an existing schema is evoked. When the evidence is consistent with the activated schema and the schema is positive, then a positive evaluation is triggered. Hence, there is a schema congruity effect (Mandler, 1982). In effect, our evaluations are driven by the degree of schema (in)congruity that we perceive when encountering a new piece of evidence. The new evidence received in conjunction with the existing schema forms the evaluation. It should be added, however, that an increasing body of literature indicates that a moderate level of incongruity between a product and a more general product category can stimulate information processing. In turn, this can lead to more favorable evaluations in relationship to products that are congruent or extremely incongruent with the schema (Meyers-Levy and Tybout, 1989).

The next section of the article briefly reviews four key theoretical areas in schema research. We then identify several areas for future research.

Four Key Theoretical Areas

Researchers have employed schema theory as the foundation for investigating a variety of applied areas of consumer behavior, including advertising, personal selling, public policy, branding, and product management. Rather than reviewing research in these applied areas, however, we have chosen to investigate contributions of consumer and psychological researchers to four key theoretical areas: (i) the effects of the level of schema congruity, (ii) the relationship of semantic memory networks to schema, (iii) the effects of moderating variables, and (iv) the impact of self-schemas on consumers.

The effects of level of schema congruity. A substantial number of studies have investigated the effects on information processing and evaluative judgments of the level of congruity between the schemas for a consumer product and its associated product category. In this case, congruity or incongruity may be defined as the extent to which the holistic configuration of attributes associated with an object and the configuration of the existing schema for the product category correspond (Mandler, 1982). In general, this research follows the logic developed by Fiske and colleagues (e.g., Fiske and Neuberg, 1990) in the social psychology literature, the work of Sujan (1985) in the consumer behavior literature, or Mandler's (1982) schema congruity hypothesis. This research suggests that when the product (or other object) matches the evoked schema, then affect transfer from the evoked category to the object occurs. In contrast, when a mismatch is perceived, more elaborate processing will be triggered. In other words, in the case of congruence (or a match), holistic processing based on the schema will take place; however, when incongruent (or a mismatch), judgments will be derived in a piecemeal fashion.

As a general statement, the literature suggests that people react positively when objects conform to expectations, and as a result are congruent with existing schema (e.g., Fiske, 1982). Mandler (1982) added to the literature, however, by proposing that objects that match the existing schema are unlikely to elicit deep cognitive processing. Therefore, the positive response generated by congruence will be weak. Conversely, when the object is incongruent with the existing schema, deeper cognitive processing is likely to be triggered. In some instances these

incongruities can be easily resolved through a process in which the individual recognizes that the object is just another example of the category. Thus, the object is recognized as a special case of an existing category (i.e., subtyping). Alternatively, the object may be recognized as residing in another category (i.e., alternative classification). Mandler (1982), and later Meyers-Levy and Tybout (1989) in the consumer behavior literature, suggest that the process of rectifying perceptions of incongruity between the object and the evoked schema will influence evaluative responses, and the process of successfully resolving the inconsistency can lead to greater positive affect. In contrast, however, extreme incongruities that cannot be easily rectified are likely to result in a negative affect. Meyers-Levy and Tybout (1989) argue that the positive effect of moderate incongruity is lacking in studies taking the Fiske (1982) approach because the effect of the process or responding to incongruity is subtle and likely to be overwhelmed when general affect toward the product or existing schema is very strong.

Again, many consumer studies take this approach. For example, Braun-LaTour, Puccinelli and Mast (2007) apply Mandler's hypothesis to investigate the effects of mood and information overload on consumer processing of ads. Their results indicate that reaction times are faster when categorizing congruent information than when categorizing incongruent information. Also, consumers in a positive mood were found to categorize faster.

The role of memory and semantic networks in schema research. Another approach adopted in the study of schemas involves the study of memory (*see* CONSUMER MEMORY PROCESSES) and semantic networks. For example, the spreading-activation theory or semantic processing proposed by Collins and Loftus (1975) is employed in social psychology and some consumer behavior research to better understand the role of memory in schemas. The spreading-activation model links concepts (or nodes) via a network. When one concept is activated, it spreads to other related concepts. Concept nodes are linked together and usually flow in both directions. Further, the links between the concept nodes will vary in "criteriality" or strength. Thus, it might be highly criterial that one concept is linked to another concept in memory (e.g., an apple node is linked to a fruit node so that an apple is recognized as a piece of fruit), but less criterial for it to be linked to another (e.g., the apple node is linked to a yellow color node). The theory suggests that when a concept node is activated, activation spreads to other nodes via the connected paths of the network. The level of activation decreases as it spreads further across the network. Thus, concepts further removed from the initial concept node that was activated will be weakly triggered while those closer are more strongly triggered. Furthermore, activation can only start at one node at a time, and activation goes away gradually or decreases over time.

Braun-LaTour and LaTour (2004) apply concepts from spreading-activation theory in an advertising study. They argue that when a consumer is exposed to an advertisement, brand-related concepts are activated. Concepts related to the spokesperson and/or the message may also be activated. Associations between the links become more automatic and stronger with repetition. Thus, a continuous ad campaign will result in better recall and recognition among consumers than a new campaign. Although not surprising, the results are consistent with schema congruity theory.

Moderation effects. Consumer researchers have identified a number of variables that moderate consumers' schema processing strategies. For example, Yoon (1997) finds that older adults exhibit greater use of schema-based processing strategies than younger adults. However, this effect is found only during nonoptimal times of the day, which for older adults is during the evening hours. In contrast, younger adults engage in detailed processing strategies regardless of the time of day.

An important finding in the schema literature is the "moderate incongruity effect" in which people prefer an option that is moderately inconsistent with a product category schema to one that is congruent with the product category schema. In two studies, the effect is found to be moderated by other variables. First, Peracchio and Tybout (1996) find that the effect occurs

to a greater extent when consumers have little prior knowledge of the product category. Thus, when consumers are knowledgeable of a product category, the "moderate incongruity effect" disappears. Campbell and Goodstein (2001) extend the research by investigating the role of perceived risk on the "moderate incongruity effect." Consistent with previous research, the authors find that moderate incongruity results in more detailed processing and evaluations that are more positive than in congruity conditions. However, perceived risk is found to moderate the effect. That is, the moderate incongruity effect is found only in low-risk product selection conditions. If perceived risk is high, then congruity (rather than moderate incongruity) leads to a more favorable evaluation.

The "moderate incongruity effect" is also investigated from the perspective of the elaboration-likelihood paradigm. In a study of ad-brand incongruency, Dahlén *et al.* (2005) find that when an ad is incongruent with a brand schema two things occur. First, the attitude toward the ad decreases. Second, brand attitude evaluations increase along with the amount of brand-related information processing. These effects are moderated, however, by the respondents' need for cognition, such that the increased brand attitude and information processing occur only among the high need for cognition respondents. Consistent with these findings is research by Lee and Thorson (2008), who find that a moderate incongruence between celebrity and product schemas increases the persuasiveness of an endorsement, but the effect is most pronounced among participants with higher levels of product involvement. Both of these studies can be interpreted from within the elaboration-likelihood theory paradigm. That is, those higher in the need for cognition and those with higher levels of product involvement engage in greater levels of information processing, which allows the respondents to cognitively handle the information in the moderate incongruency conditions.

In an interesting study, Aggarwal and McGill (2007) investigate the role of schemas in evaluating anthropomorphized products. The results of their first two studies reveal that the greater the extent that a product matches the schema for what a human looks like, the more positive the evaluation of the product when the text of the ad anthropomorphizes the product. For example, in one of the experiments an automobile is depicted as talking to the consumer when the ad begins "Hi, I am a Lexus." A third study, however, shows that the findings depend upon whether the affective tag associated with the human schema is positive. Thus, if the human schema that is activated is negative, the evaluation of the brand becomes more negative. In the study, beverage bottles are shown in pairs. In the copy of the ad, the bottles are described as either good twins, who will make their parents proud, or as evil twins who will conquer your city. The results reveal that the affective tag of "good twins" versus "evil twins" moderates the results. Thus, the evaluation of the brand is enhanced in the "good twin" condition in comparison to the "evil twin" condition. This study is important because it shows that schema congruity can increase or decrease evaluations depending upon the affective tag associated with the schema that is evoked.

The role of self-schemas in consumer behavior. Aaker's (1999) research on the effect of self-schemas on consumer attitudes toward a brand provides the foundation for additional work on self-schemas in a branding context. Aaker proposes that the self (*see* POSSESSIONS AND SELF) is malleable and that one's self-schema is influenced by personality and situational factors. Various self-conceptions (i.e., self-schema) may be activated or deactivated depending on how salient they are, which can be influenced by the existing social situation or by the traits that are evoked by a particular experience or memory. In her research she focuses on those traits that are deemed to be schematic because they are at that moment descriptive or important to the individual. Consistent with social psychology research, Aaker argues that people maintain a self-schema because they desire consistency and positivity. When presented with information that is inconsistent with the self-schema, then individuals are less likely to accurately recall the information. Further, negative affect is more likely to result if self-relevant information is inconsistent with the self-schema.

Aaker (1999) proposes that a brand (and the set of personality traits that the brand embodies) often serves a "self-expressive" purpose for consumers. As consumers, we select and utilize brands that allow us to make a certain impression or that best represent us at a given point in time. As a result, consumers who are schematic on a particular personality trait should demonstrate a preference for brands that are also highly descriptive on that trait. Aaker refers to this as "self-congruity." While consumers' schematic traits typically guide their behavior, she also acknowledges that sometimes consumers' behaviors express who they wish to be (desired self) or strive to be (ideal self) rather than the actual self. Thus, each individual holds a variety of self-conceptions. She also suggests that depending on the current social situation, certain situational cues can become salient and influence the conceptualization of the self so that one trait self-schema is more acceptable than others. On the basis of these situational cues, people will tailor their behavior to match the trait that is primed by the situation (i.e., "situation congruity" takes place).

Batra and Homer (2004) adopt Aaker's schematicity framework in their research on brand image beliefs. They find that personality associations of celebrity endorsers serve to reinforce consumer beliefs about the brand. That is, if the spokesperson is perceived to be fun or sophisticated, then the brand is likely to be perceived to be more fun or more sophisticated. For brand image beliefs to be evoked, the brand's schema needs to fit the consumer's schema. Consistent with these findings, Worth, Smith and Mackie (1992) find that brands are evaluated more favorably if the description of the brand matches the consumer's self-schema. More specifically, these authors found that the gender schematicity of the respondents is assessed in terms of its degree of masculinity versus femininity. Respondents prefer products whose description matches their self-perceived schema for masculinity or femininity. Similar findings are found for self-schema based upon the separateness–connectedness of the subjects (Wang and Mowen, 1997). Thus, respondents who rate themselves as more connected to others preferred products described in terms of their ability to bring people together. In contrast, respondents who rate themselves as more separated prefer products described in terms of their ability to help maintain the person's individuality.

Directions for Future Research

One direction for future research involves the investigation of the nature of self-schemas. In the development of the 3M Model of Motivation and Personality, Mowen (2000) and his colleagues (e.g., Mowen and Sujan, 2005) propose a control theory model in which a system of personality traits provides a schema for interpreting stimuli and outcomes as well as for influencing attitudes and behavior. Thus, traits do not operate in isolation of each other. Rather, in conjunction with situational forces, multiple traits combine to form a schema that influences attitudes and behaviors. Importantly, past research has tended to investigate how single personality traits (e.g., need for cognition or extroversion) influence evaluations if they are congruent or incongruent with the description of a product. For example, as noted earlier, masculinity–femininity and separateness–connectedness can be conceptualized as traits and have been shown to positively impact brand evaluations if appropriately matched to a product's description. To the present authors' knowledge, however, researchers have not investigated multiple traits acting together in experimental studies.

How can an advertisement, brand, or endorser simultaneously activate a schema associated with multiple traits? Consider the development of a television advertisement. These ads will have multiple elements, including a message, music, a situational context, and perhaps an endorser. Each ad component can be used to activate a different personality trait. For example, the message could be composed of high-quality arguments, which would appeal to an individual high in the need for cognition. Similarly, the music could be based upon a song that emphasizes togetherness (e.g., "Love Will Keep Us Together"), which would activate a connectedness trait. The situation could be composed of a party atmosphere in which an endorser who is shown to be highly gregarious, which would activate an extroversion trait. These ideas suggest that a critical aspect of developing TV ads is

the art of combining music, copy, situation, and endorser so as to effectively activate schema that are maximally congruent the product's schema as well as the self-schema of the target audience. These ideas also illustrate the crudeness of present schema research, which looks at only one trait at a time.

The idea that multiple traits act to form a self-schema suggests the possibility of bringing science into the art of creating advertisements. We believe that the complex interaction among a multitrait person schema, a product schema, and a situational context is why the development of advertisements can be described as an art. One possibility for incorporating a greater level of science into the art, however, is by identifying the trait antecedents of the specific behavior that ad developers hope to influence. In the hierarchical structured 3M Model, these highly specific behavioral dispositions are called *surface traits*, which emerge in part from combination of deeper level traits, including situational, compound, and elemental traits.

Here is an example. Suppose that the advertiser is marketing a product that is positioned as the leader in its field. The goal is to attach a leadership schema to the product. This schema would be based upon traits that are associated with individuals having a strong leadership propensity. The goal would be to create an advertisement that portrays a product schema that is congruent with a leadership self-schema. Research by Flaherty *et al.* (2009) identifies the system of traits of individuals with high levels of leadership propensity. This self-schema includes the deeper level traits of competitiveness, self-efficacy, and the need for material resources. With this information it would be straightforward to create an advertisement that employs an endorser who is self-confident (i.e., indicating self-efficacy), that takes place in a luxurious setting (suggesting materialism), and that includes music suggesting competition (e.g., the song "Eye of the Tiger"). Schema theory suggests that such a high level of congruity between the product's schema and self-schema of the target market should enhance the effectiveness of the advertisement. Future research, however, is required to test these ideas.

A second arena for future research involves integrating the research on schema congruity with the literature on fit and the match-up hypothesis. For example, Till and Busler (2000) propose that a match-up occurs when there is a fit between the characteristics of the product spokesperson and the product that is endorsed. That is, when fit, match-up, or similarity occurs between two entities, there is a greater likelihood of the transfer of knowledge and affect from one stimulus to another. This idea is consistent with both spreading-activation theory and balance theory, which places the literature squarely within the domain of schema congruity.

The concepts of fit and match-up have been applied to numerous cases in which congruence is sought in the characteristics between two entities. Across the studies, however, the authors are inconsistent in whether they cite the literature on schema congruity. The following are examples of the topics of studies investigating stimulus congruity from the perspective of fit and/or match-up.

- Fit of endorser with product
- Fit of country-of-origin with product
- Fit of style of music with brand image of retail store
- Fit of music with advertising message
- Fit of product extension with brand
- Fit of brand image with social cause in cause-related marketing
- Fit of advertisement with TV program
- Fit of product with TV program
- Fit of product with an ambient scent/smell
- Fit of attractiveness of service provider with nature of the service.

Each of the pairs of entities can be investigated from the perspective of schema theory. As such, the two literatures should be integrated. Interestingly, while the authors of the studies focusing on match-up/fit will sometimes cite schema theory, the authors investigating phenomena from a schema congruity perspective far less frequently cite the literature on match-up/fit.

The present authors believe that a key goal for future research is the development of a general model that could be possibly called the "schema congruency theory." The model will require the development of an enhanced version of schema theory that is able to explain

the effects of moderating variables as well as handle the multifaceted nature of schema. The new theory of schema congruence would then be able to explain and predict consumer reactions to the level of congruency, which includes issues of match-up/fit among stimuli, whether product-based, self-based, other person-based, and/or situation-based.

Bibliography

Aaker, J.L. (1999) The malleable self: the role of self-expression in persuasion. *Journal of Marketing Research*, **36**, 45–57.

Aggarwal, P. and McGill, A.L. (2007) Is that care smiling at me? Schema congruity as a basis for evaluating anthropomorphized products. *Journal of Consumer Research*, **34**, 468–479.

Anderson, S.M. and Cole, S.W. (1990) Do I know you? The role of significant others in general social perception. *Journal of Personality and Social Psychology*, **59**, 384–399.

Asch, S.E. (1946) Forming impressions of personality. *Journal of Abnormal and Social Psychology*, **4**, 1230–1240.

Batra, R. and Homer, P.M. (2004) The situational impact of brand image beliefs. *Journal of Consumer Psychology*, **14**, 318–330.

Braun, K.A. (1999) Post-experience effects on consumer memory. *Journal of Consumer Research*, **25**, 319–334.

Braun-LaTour, K.A. and LaTour, M.S. (2004) Assessing the long-term impact of a consistent advertising campaign on consumer memory. *Journal of Advertising*, **33**, 49–61.

Braun-LaTour, K.A., Puccinelli, N.M. and Mast, F.W. (2007) Mood, information congruency, and overload. *Journal of Business Research*, **60**, 1109–1116.

Cacioppo, J.T., Petty, R.E. and Sidera, J.A. (1982) The effects of a salient self-schema on the evaluation of proattitudinal editorials: Top-down versus bottom-up message processing. *Journal of Experimental Social Psychology*, **18**, 324–338.

Campbell, M.C. and Goodstein, R.C. (2001) The moderating effect of perceived risk on consumers' evaluations of product incongruity: preference for the norm. *Journal of Consumer Research*, **28**, 439–449.

Cantor, N. and Mischel, W. (1979) Prototypes in person perception, in *Advances in Experimental Social Psychology*, vol. 12 (ed. L. Berkowitz.), Academic Press, New York, pp. 3–52.

Collins, A.M. and Loftus, E.F. (1975) A spreading-activation theory of semantic processing. *Psychological Review*, **82**, 407–428.

Dahlén, M., Lange, F., Sjödin, H. and Törn, F. (2005) Effects of ad-brand incongruency. *Journal of Current Issues*, **27**, 1–12.

Fiske, S.T. (1982) Schema-triggered affect: applications to social perception, in *Affect and Cognition: The Seventeenth Annual Carnegie Symposium on Cognition* (eds M.S. Clark and S.T. Fiske), Lawrence Erlbaum, Hillsdale, NJ, pp. 55–78.

Fiske, S.T. and Neuberg, S.L. (1990) A continuum of impression formation, from category-based to individuating processes: influences of information and motivation on attention and interpretation, in *Advances in Experimental Social Psychology*, vol. 2 (ed. M.P. Zanna), Academic Press, New York, pp. 1–74.

Fiske, S.T. and Taylor, S.E. (1991) *Social Cognition*, 2nd edn, McGraw-Hill, New York.

Flaherty, E.T., Mowen, J.C., Brown, T.J. and Marshall, G.W. (2009) Leadership propensity and sales performance among sales personnel and managers in a specialty retail store setting. *Journal of Personal Selling & Sales Management*, **39**, 39–55.

Hastie, R. (1980) Memory for behavioral information that confirms or contradicts a personality impression, in *Person Memory: The Cognitive Basis of Social Perception* (ed. R. Hastie *et al.*), Erlbaum, Hillsdale, NJ, 155–177.

Heider, F. (1958) *The Psychology of Interpersonal Relations*, Wiley, New York.

Kahneman, D. (1973) *Attention and Effort*, Prentice-Hall, Englewood Cliffs, NJ.

Kanungo, R.N. and Pang, S. (1973) Effects of human models on perceived product quality. *Journal of Applied Psychology*, **57**, 172–178.

Kihlstrom, J.F., Cantor, N., Albright, J.S. *et al.* (1988) Information processing and the study of the self, in *Advances in Experimental Social Psychology*, (ed. L. Berkowitz), vol. 21, Academic Press, New York, pp. 145–180.

Lee, J.G. and Thorson, E. (2008) The impact of celebrity-product incongruence on the effectiveness of product endorsement. *Journal of Advertising Research*, **48**, 433–449.

Lynch, J. and Schuler, D. (1994) The matchup effect of spokesperson and product congruency: a schema theory interpretation. *Psychology & Marketing*, **11**, 417–445.

Macrae, C.N. and Bodenhausen, G.V. (2000) Social cognition: thinking categorically about others. *Annual Review of Psychology*, **51**, 93–120.

Mandler, G. (1982) The structure of value: accounting for taste, in *Affect and Cognition: The Seventeenth Annual Carnegie Symposium on Cognition* (eds M.S. Clark and S.T. Fiske), Erlbaum, Hillsdale, NJ, pp. 3–36.

Markus, H. (1977) Self-schema and processing information about the self. *Journal of Personality and Social Psychology*, **35**, 63–78.

Markus, H. and Wurf, E. (1987) The dynamic self-concept: a social psychological perspective, in *Annual Review of Psychology*, vol. 38 (ed. M.R. Rosenzweig and W. Porter), Annual Reviews Inc., Palo Alto, CA, pp. 299–337.

Meyers-Levy, J. and Tybout, A.M. (1989) Schema congruity as a basis for product evaluation. *Journal of Consumer Research*, **16**, 39–54.

Mowen, J.C. (2000) *The 3M Model of Motivation and Personality: Theory and Empirical Applications to Consumer Behavior*, Kluwer Academic Press, Boston, MA.

Mowen, J.C. and Sujan, H. (2005) Volunteer behavior: a hierarchical model approach for investigating its trait and functional motive antecedents. *Journal of Consumer Psychology*, **15**, 170–182.

Peracchio, L.A. and Tybout, A.M. (1996) The moderating role of prior knowledge in schema-based product evaluation. *Journal of Consumer Research*, **23**, 177–192.

Sujan, M. (1985) Consumer knowledge: effects on evaluation strategies mediating consumer judgments. *Journal of Consumer Research*, **12**, 31–45.

Sujan, M. and Bettman, J.R. (1989) The effects of brand positioning strategies on consumers' brand and category perceptions: Some insights from schema research. *Journal of Consumer Research*, **26**, 454–467.

Sujan, M., Bettman, J.R. and Sujan, H. (1986) Effects of consumer expectations on information processing in selling encounters. *Journal of Marketing Research*, **23**, 346–353.

Taylor, S.E. and Fiske, S.T. (1978) Salience, attention and attribution: top of the head phenomenon, in *Advances in Experimental Social Psychology*, vol. 11 (ed. L. Berkowitz's), Academic Press, New York, pp. 249–288.

Till, B.D. and Busler, M. (2000) The match-up hypothesis: physical attractiveness, expertise, and the role of fit on brand attitude, purchase intent, and brand beliefs. *Journal of Advertising*, **29**, 1–13.

Unnava, H.R. and Burnkrant, R.E. (1991) An imagery-processing view of the role of pictures in print advertisements. *Journal of Marketing Research*, **28**, 226–231.

Wang, C.L. and Mowen, J.C. (1997) The separateness-connectedness self-schema: scale development and application to message construction. *Psychology & Marketing*, **14**, 185–207.

Wheeler, S.C., Petty, R.E. and Biser, G.Y. (2005) Self-schema matching and attitude change: situational and dispositional determinants of message elaboration. *Journal of Consumer Research*, **31**, 787–797.

Worth, L.T., Smith, J. and Mackie, D.M. (1992) Gender schematicity and preference for gender-types products. *Psychology & Marketing*, **9**, 17–30.

Yoon, C. (1997) Age differences in consumers' processing strategies: an investigation of moderating influences. *Journal of Consumer Research*, **24**, 329–342.

consumer intentions

Richard P. Bagozzi

An *intention* has been defined as a *person's commitment, plan, or decision to carry out an action or achieve a goal* (Eagly and Chaiken, 1993), and in fact has been used synonymously at times with choice, decision, and plan. All these usages more generally fall under the label *volition*. Psychologist Ajzen (1991, p. 181) conceives of intentions rather broadly as "indicators of how hard people are willing to try, of how much of an effort they are planning to exert." This definition seems too broad in that it encompasses (i) motivation, which is better construed as an antecedent of intention and (ii) planning, which constitutes a mental activity or process that often occurs after one forms an intention to pursue a goal or perform an action.

The need for a narrower definition of intention can be seen in Lewin's (1951, pp. 95–96) specification of the role of volition in action: "[A] complete intentional action is conceived as follows: Its first phase is a motivational process, whether a brief or a protected vigorous struggle of motives; the second phase is an [mental] act of choice, decision, or intention, terminating this struggle; the third phase is the consummatory intentional action itself." This clearly differentiates intention from motivation and action and situates it between these concepts: motivation→intention→action.

The most common type of intention is the personal or I-intention to pursue a goal or perform an action by oneself. Notice that one can have an intention to pursue a goal, accomplish an end, or produce an outcome ("I intend to lose body weight") or an intention to execute an act ("I intend to buy a new LED television"). Both goal intentions and action intentions can be expressed noncontingently, as phased above, or contingently (e.g., "I intend to buy a new LED television, as soon as the price drops below $3000").

Gollwitzer proposed an important kind of intention, which he termed, *implementation intentions* (e.g., Gollwitzer and Brandstätter, 1997). These involve planning when, where, and how to act. Typically, a gap in time, often significant, exists between intention formation and behavioral execution. Implementation intentions serve cognitively to provide mental representations of opportunities to act and volitionally "create strong mental links between intended situations and behavior" and "in the presence of the critical situation, the intended behavior will be elicited automatically" (Gollwitzer and Brandstätter, 1997, p. 196). Thus one's intention to buy milk at the end of the day, made in the morning before work begins, may be recalled by the sight of a favorite store that one passes while returning home from work and enacted straight away. Because intentions frequently form first and lead to planning, many researchers might prefer to term "implementation intentions," *planning*, to allow for separate mental states and processes in this regard.

Intentions might lead directly to action straight away or after a gap in time. But they also have been shown to differ in degree of well-formedness and to be moderated in their effects by level of effort required in action execution and by correspondence or independence between goal commitment and planning. Further, various mediators may occur between intentions and behavior, such as the need to deal with temptations, impediments, weakness-of-will, and monitoring progress in goal pursuit. For a review of moderators and mediators of the intention-action linkage, *see* Bagozzi (2006a, p. 19ff).

A common role for intentions in so-called goal-directed action is to bridge desires and their downstream effects. The consumer core captures this role of intentions and can be expressed as follows: goal desire→goal intention→action desire→action intention→action (see Bagozzi, 2006b). An important way that self-regulation occurs is in the management of the desire-to-intention links sketched above (see Bagozzi, 2006a, 2006b). Here desires are modulated through the imposition of personal standards of conduct, ethics, or moral imperatives. (*see also* SELF-REGULATION; CONSUMER DESIRE).

Recently the meaning of intentions has been expanded to include shared or collective intentions (Bagozzi, 2000, 2005). Thus, for example, persons in an intimate relationship might speak of "our intention to see Tchaikovsky's Swan

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Lake"; a football player may mention "the team's plan to implement a new defensive scheme"; corporation spokespersons might announce "the firm's hostile intention to take over another firm"; and the President of a country might remark that "our people intend to eliminate poverty by 2020." These examples – referring, respectively, to a two-person dyad, a small group, an organization, and a collectivity – illustrate that people often use social notions of intentions in ordinary speech, whether referring to informal or formal groups.

There are two types of collective intentions. One is a personal intention to do something with a group of people or contribute to, or do one's part of a group activity (e.g., "I intend to prepare the holiday dinner with my sisters"). Notice that a person can have an intention to act as an actor, yet the action can be self-construed as an individual act performed alone (an I-intention to do an individual act by oneself) or as a member of a group (an I-intention to do one's part of a group act). The latter is one type of collective action.

A qualitatively different form of a collective intention is what might be called a *we-intention*. A we-intention is a collective intention rooted in a person's self-conception as a member of a particular group (e.g., an organization or institution) or social category (e.g., one's gender or one's ethnicity), and action is conceived as either the group acting or the person acting as an agent of, or with, the group. A we-intention can be expressed in two forms: a shared we-intention articulated in the form of "I intend that our group/we act" (e.g., "I intend that our family visit Sea World, San Diego, next vacation") or a communal we-intention explicated in the form, "We (i.e., "I and the group to which I belong") intend to act" (e.g., "We intend to sponsor an exchange student in our home next year"). Collective intentions open up new avenues for exploring consumer behavior and are grounded in plural subject theory (Bagozzi, 2000, 2005).

Traditionally, such psychological determinants of behavior as attitudes, motives, emotions, values, felt social pressure, self-efficacy, and perceived behavioral control have been conceived to work through intentions to influence behavior. More recently desires have been considered to be proximal causes of intentions, channeling the effects of multiple psychological determinants (see consumer desires in this volume).

Bibliography

Ajzen, I. (1991) The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, **50**, 179–211.

Bagozzi, R.P. (2000) On the concept of intentional social action in consumer behavior. *Journal of Consumer Research*, **27**, 388–396.

Bagozzi, R.P. (2005) Socializing marketing. *Marketing: Journal of Research in Management*, **1**, 101–110.

Bagozzi, R.P. (2006a) Consumer action: automaticity, purposiveness, and self-regulation, in *Review of Marketing Research* (ed. N.K. Malhotra) Sharpe, Armonk, pp. 3–42.

Bagozzi, R.P. (2006b) Explaining consumer behavior and consumer action: from fragmentation to unity. *Seoul Journal of Business*, **12**, 111–143.

Eagly, A.H. and Chaiken, S. (1993) *The Psychology of Attitudes*, Harcourt Brace Jovanovich, Fort Worth.

Lewin, K. (1951) *Field theory in social science*, Harper and Row, New York.

Gollwitzer, P.M. and Brandstätter, V. (1997) Implementation intentions and effective goal pursuit. *Journal of Personality and Social Psychology*, **73**, 186–199.

consumer behavior across literacy and resource barriers

Madhu Viswanathan

With few exceptions (Hill, 1991; Alwitt, 1995; Arnould and Mohr, 2005), much of what is known about consumer behavior has come from studying relatively resource-rich, literate consumers, usually in advanced economies. This article focuses on low-literate, low-income consumers in the United States and subsistence consumers in India, that is, those living in circumstances of widespread poverty. Each of these groups is described and compared and contrasted with what is known about consumer behavior in more conventional settings. Thus, the focus is on consumer behavior across literacy and resource barriers. Implications for marketing in general are also briefly discussed.

Low-literate Low-income Consumer Behavior in The United States

Background. Literacy relates to reading and writing skills, whereas numeracy relates to counting. Functional literacy relates to possessing the reading, writing, and counting skills to function in day-to-day life (Kirsch and Guthrie, 1997). Literacy and functional literacy are used interchangeably in this article. Low literacy, of course, is associated with low income as well, and the discussion emphasizes low-literate, low-income consumers in the United States, disentangling these two factors where possible. Literacy rates in the United States, until a few decades ago, were reported to be as high as 99%. However, such statistics were based on a question in the Census as to whether an individual was literate. Subsequently, when grade-equivalent tasks were used (e.g., if you are in a certain grade, you should be able to write a check or read a package label), the estimates of low literacy have varied from 20% to considerably higher. Statistics about literacy carry a sense of prestige with them and therefore, often reflect biases at individual and more aggregate state or country levels. The 2002 National Assessment of Adult Literacy (NAAL) showed that at least 22% of US consumers lack skills to perform retail tasks (e.g., calculating unit prices and price discounts, comparing product attributes), and between 34% and 55% lack skills to look up reference materials to identify foods containing a particular vitamin (Kutner, Greenberg, and Baer, 2005).

Method. In describing literate versus low-literate consumers, it is important to understand that this is a continuum and that there is no demarcation that makes someone nonliterate, hence the preference here to use the term, low-literate. Also relevant here is how low-literate consumers can be studied. Experiments and surveys are not the most effective approaches, unless very carefully designed (Viswanathan, Gau, and Chaturvedi, 2008a). In this work, observations of classrooms at adult education centers, observations of shopping trips, and interviews of teachers and students were used. Typically, such centers have students divided by grade-equivalent levels, such as 0–4, 5–8, and 9–12, based on reading and math tests administered periodically, that is, ranging from low to moderate levels of literacy. Although a number of insights are offered below, it should be noted that some of them apply more to those at the 0–4 level than say those at the 5–8 level. Both vulnerabilities and strengths of low-literate consumers were observed in this work and that is important to acknowledge as well. Low-literate consumers overcome constraints borne out of their circumstances, sometimes in ingenious ways, and display extraordinary resilience in negotiating a marketplace that assumes a certain level of literacy. Over time, surveys and experiments have been employed as well.

Illustrative findings. In observations of low-literate consumers while shopping, most striking were the things that literate consumers may take for granted. A typical trip to the grocery store for a literate consumer may involve some trade-offs between price and attributes of products. If the intention is a cash purchase and one is short by a few dollars, it is usually not a big issue to leave behind some items and attribute things to one's forgetfulness. But this typical trip does not begin to describe the kinds of issues that take up much of the effort of low-literate shoppers, particularly those at the low end of the continuum.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Locating a product can take considerable time, when reading signs is difficult or just not possible. Therefore, once a product is located, it may be bought. It is not always easy to be asking store employees for directions, particularly when an individual has spent a lifetime with low literacy and it has been cause for many travails. Identifying the correct volume can be effortful – finding 150 candles involves finding two different packets of a 100 and a 50. With the many different price tags and sales signs, finding the bottom-line price can be very confusing. Also, low-literate individuals may need to rely on written computations to find out the price of multiple units when the price of one is known. "% off"sales signs can be another cause for concern as some may completely avoid them owing to difficulty computing the final price or embarrassment with having to ask a store employee, whereas others may use them sometimes, say, when 50% off, as computing half off may be relatively easier. Computing the total on a shopping trip may be difficult, leading to solutions such as moving currency bills after each item is added to a cart or visually subtracting from an imagined set of currency bills. Similarly, allowing for taxes can be a significant concern leading to rules of thumb like buying one thing if one has five dollars. Unit prices are essentially abstractions that are difficult to understand and often ignored. Similarly, nutrient values may not be used and "% DVs," serving sizes, and even expiry dates misunderstood or misused.

Adding to cognitive difficulties, issues such as being short of money at the checkout counter are cause for despair, as this is a result of one's low literacy rather than forgetfulness. Similarly, having enough money can be cause for celebration. Underlying their shopping interactions for low-literate consumers is the need to maintain self-esteem and avoid being exposed for their low literacy. Thus, in addition to cognitive issues, emotional issues are also central, with anxiety being a common aspect of seemingly mundane shopping activities.

Cognitive tendencies. Some of these findings with cognitive tendencies, decision-making aspects, and coping strategies are summarized. A tendency that is most striking is one of concrete thinking, focusing on single pieces of information such as price, without abstracting across, say price and size or price and other attributes (Viswanathan, Rosa, and Harris, 2005; Gau and Viswanathan, 2008). Buying the cheapest product without attention to size is an example of this tendency. It manifests in a number of ways, through seeking familiar stores, learning to use expiry date or other concrete information, such as numerical information, without understanding its meaning. In a study of low-literate peasants in Central Asia in the early 1900s, a Russian psychologist asked participants to view tools (e.g., an axe, a hatchet, a log, and a saw) and asked them to group three of these objects that can be described by a word or belong together (Luria, 1976). Participants responded in terms of how they could chop firewood to stay warm, that is, how they could use what they saw in day-to-day life. Thus, low literacy and difficulty with abstractions tends to lead consumers to think in the immediate, visual, graphic, here-and-now world of how they can use things.

Another cognitive tendency is pictographic thinking (Viswanathan *et al.*, 2005; Gau and Viswanathan, 2008). Although consumers in general depend on pictures, pictographic thinking is qualitatively different. It means viewing text such as brand names not as something to read but as an image, and thus remembering brands to buy through pictorial elements. Pictographic thinking also manifests in visualizing quantities to buy rather than using symbolic information. For instance, when buying an ingredient like sugar to bake a cake, low-literate consumers may visualize baking the cake and pouring sugar into it and buy the package of corresponding size. Pictographic thinking may even involve "counting" by visualizing currency bills. Pattern-matching prescriptions with medicine packages to find the right medicine is another approach. Whereas low-literate consumers may have poorer memory for textual information when compared to consumers with higher literacy, their memory for pictorial information, such as brand signatures, can be as good (Viswanathan *et al.*, 2009). Such improvement in memory appears to arise from familiar pictorial elements.

Decision making, emotional trade-offs, and coping strategies. Decision rules may range from single-attribute decisions (buy the cheapest) to habitual decision making or even random decision making (picking up the first item in a product category that is seen, for a planned or unplanned purchase). As noted earlier, sometimes, locating a product can lead to its purchase. This should be contrasted with conventional decision-making models, which cover a number of steps. Also intertwined with decisions are emotional trade-offs. For instance, low-literate consumers may trade off utility to save embarrassment, as maintaining self-esteem is central when they negotiate the marketplace. Having sufficient money at the counter may be cause for celebration, whereas being "caught" short may be cause for despair. The anxiety associated with negotiating a complex marketplace while lacking in literacy and trying to avoid being exposed on this account is the larger emotional context where day-to-day shopping occurs. Along similar lines, Adkins and Ozanne (2005) identify types of low-literate consumers on the basis of dimensions such as acceptance or rejection of stigma arising from low literacy. They find a number of coping strategies including avoidance, dependence on others, self-esteem maintenance, and social deception.

Low-literate consumers cope in a number of ways – depending on others, coming up with rudimentary defensive rules like buying one item from the menu at a time to avoid being short of money, and even giving all the money they have at the checkout counter. Some of the ways in which they cope are ingenious, such as in "counting" by visualizing currency bills. They often depend on others, sometimes even helpful store employees, to help them with shopping and avoid situations and stores that are likely to lead to embarrassment, striving to maintain self-esteem in their interactions.

Low-literate shopping behavior. A model of low-literate shopping behavior is presented in Figure 1 purely for illustrative purposes to contrast low-literate consumer behavior with more conventional consumer behavior. It should be noted that this model captures consumer behavior especially for 0-4 level adult education students, setting up a sharp contrast with consumers with higher levels of literacy. In terms of generic models of decision making, low-literate consumers may bypass or spend minimal effort in steps such as prepurchase search and evaluation of alternatives. Alternatively, consumers bypass the entire decision process by mimicking others or by delegating responsibility to others. As indicated in the direct link from "need recognition" to "purchase," some consumers may completely delegate decision making to others, highlighting the theme of dependence on others. The first step in the decision-making process is location of a product. Owing to the amount of effort expended in locating a product in the store, location may be followed by purchase. Location of a product may also be followed by obtaining the correct number of units of a product. Alternatively, the next step may be to expend effort in looking for a price, a distinct step in the process. The process may be "short-circuited" at this point by either reshelving or buying the product caused by lack of success in finding the price. The next step may be reading a price and iteratively looking for a price owing to confusion in identifying a price tag or interpreting multiple price stickers. Again, the process may be short-circuited through reshelving or purchase. The next step may be to locate the correct volume or number of units needed. In instances where a combination of package sizes are needed, significant effort may be needed in locating the correct packages. Location of needed volume may precede location of a price and reading a price, and may follow location of a product and may immediately lead to purchase. Location of volume may also be followed by looking for a price and/or reading a price. The next step is to compute a total price with some iteration needed here to perform the computations. Iterations may also be required here in looking for the price again or rereading the price, or identifying the correct volume. Reshelving or purchase culminates the decision-making sequence.

Although not strictly comparable to generic models of decision making and information processing adapted from the literature (Bettman, 1979), the shopping behavior model for functionally low-literate consumers highlights the importance of one or two stages in other models.

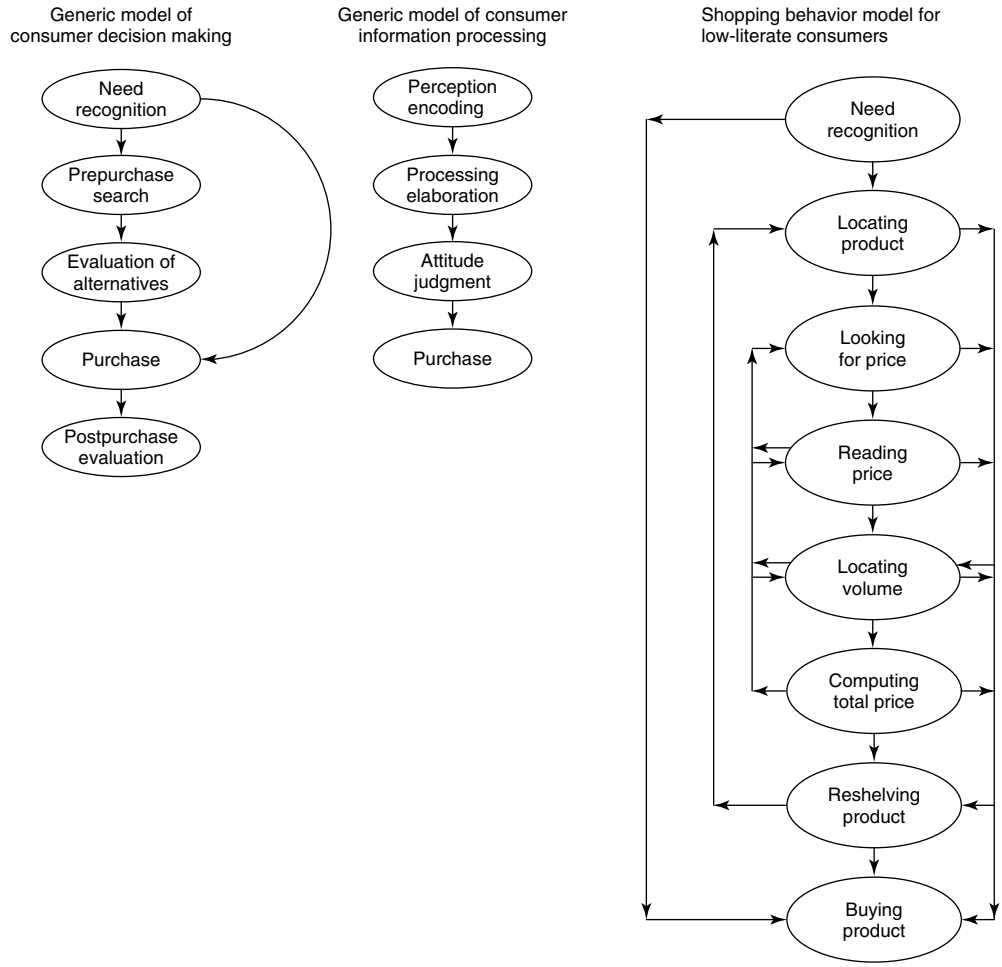


Figure 1 An illustration of shopping behavior for low-literate consumers.

In contrast to these typical decision-making and information processing models, the emphasis here is on perception of information, leading to surface-level processing of primarily one or very few pieces of product information (e.g., price information), often with error. Whereas typical information processing models include perception, encoding, elaboration, evaluation, and then purchase (Bettman, 1979), the emphasis for low-literate consumers is on perception with regard to location of product and identification of price. Encoding may be of a transient nature, with a view toward completing the immediate purchase task rather than leading to memory and integration with prior knowledge.

Importantly, whereas this model may resemble low-involvement models of decision making and models of repeat purchase in terms of lack of active information seeking and lack of comparison along attributes (cf., Zaichkowsky, 1985), fundamental processing differences are noteworthy. The decision-making process involves bypassing of several steps when compared to traditional decision-making models. Rather than arising from a lack of motivation that characterizes low-involvement decision making and repeat purchases, this process can be quite

effortful. Thus, striking differences among consumers who differ in levels of literacy are illustrated by contrasting decision-making processes.

Subsistence Consumer Behavior

Background. The discussion now moves further across the resource spectrum to examine subsistence marketplaces and consumer behavior. Research in poverty contexts in India provides the foundation for the discussion. When compared to the United States, such contexts are characterized by more widespread poverty and higher rates of low literacy as well. Much of humanity lives on the equivalent of two to four dollars a day (Prahalad, 2005). However, much of what is known about consumer behavior has been based on studying relatively resource-rich, literate consumers, mostly in advanced economies.

Method. The work here was in low-income urban neighborhoods and rural areas in South India, using a range of methods including in-depth interviews of buyers and sellers (Viswanathan, 2007; Viswanathan, Gajendiran, and Venkatesan, 2008; Viswanathan, Rosa, and Ruth, 2010). Such subsistence communities are characterized by deprivation on many fronts. Uncertainty pervades every facet of life as basic infrastructure is either unreliable or absent. For example, rural areas are characterized by unreliable transportation or electricity. Central consumption events like cooking a staple, say rice, may be dependent on many factors – the quality of rice and water, availability of cooking fuel, and so on. There are, of course, many differences between urban and rural areas as well, with the latter having lack of access to markets and shops and more severe infrastructure problems.

Illustrative findings. In such marketplaces, subsistence consumers often have relationships with a single store, such as a neighborhood retail store that is conveniently located and charges higher prices but allows purchase on credit in times of need, or a larger reseller. Through such relationships, often with the neighborhood retail store owner, subsistence consumers multiply the value of their patronage for a small seller and expect responsive customer service in return. It may be noted here that an understanding of relatively resource-rich settings may lead to misconceptions about subsistence consumer behavior. For instance, it may seem that buying from larger resellers at the beginning of the month and stocking up is the rational course of action. However, this implicitly assumes a stable income and the lack of frequent crises, certainties that are beyond the reach of subsistence consumers. Fear of crises necessitates buying on credit from the neighborhood retailer in exchange for consumer loyalty. Also interesting here is how the longer-term relationship rather than any transaction, and particularly an impersonal one, takes precedence as subsistence consumers and sellers leverage relationships to multiply the economic value and leverage social relations to compensate for the lack of marketplace infrastructure prevalent in resource-rich settings.

Subsistence consumers have to focus on products that serve immediate and basic needs, such as food and clothing, and emergency medical care. Their choices may be between buying a product, making it (e.g., indigenous soap substitute), or, of course, forgoing it. However, they aspire for quality brands that serve their needs and also for products such as cell phones that greatly improve their ability to communicate and can serve vital needs (e.g., for emergencies). They also aspire for a better future, making sacrifices to obtain better education or health care for their children – a driving motivation in many parents' lives. Subsistence consumers are also very adaptive and willing to experiment with new technology, such as cell phones, that serve vital communication needs. Their adaptivity is borne out by the need to survive and the experience of coping with extreme uncertainty in many realms of life that are taken for granted in relatively affluent settings. Thus, to fully understand subsistence consumer behavior, preconceptions that they require only some basic products or that products in demand would have to be of low technology need to be dispelled. Similarly, quality is very central, as scarce resources have to be used for products that will better subsistence consumers' immediate and basic life circumstances. Otherwise, subsistence consumers have

the resilience to forgo products or the resourcefulness to make them.

The issues discussed for low literacy in the United States apply in subsistence marketplaces as well, in terms of concrete thinking and pictographic thinking. They engage in concrete thinking, such as in fixating on concrete numbers denoting price, such as maximum retail price, a requirement in the Indian marketplace. However, they may do so without understanding what it means. This is akin to low-literate consumers in the United States who may depend on expiry date without understanding its meaning. Pictographic thinking is also pervasive, such as in pattern-matching prescriptions with medicine labels. Self-esteem issues are also relevant, and although both poverty and low literacy is pervasive, the stigma attached to the latter is perhaps more acute, with lack of education often being attributed with the lack of ability to overcome poverty. Low literacy can lead to fear of conversations in shopping contexts, and feelings of futility about arguing with shopkeepers or questioning quality.

One-to-one interactional marketplaces. In a larger context of widespread and extreme poverty and low literacy, subsistence consumers negotiate the marketplace in ways that require face-to-face interactions and oral communication in the native language. Such communication does not require specific basic literacy skills, but rather knowledge of spoken language that individuals acquire growing up and social skills. Subsistence consumers can learn marketplace skills through conversations and observations on streets and other public locations. Conversations with sellers can lead to development of skills and self-confidence. Moreover, consumers themselves may often be vendors as a way to subsist and survive, thus learning valuable skills that transfer to their consumer behavior, such as negotiating, counting, and so forth. What is interesting here is the widespread existence of distinct markets in terms of the poor and the higher strata. Clothes can be purchased at department stores or on the sidewalk for diametrically opposite price points.

In this one-on-one interactional marketplace, subsistence consumers may be resource poor but network rich. With an opportunity to participate through oral communications and observations, such a marketplace can be a stepping stone to gaining important skills, awareness of rights, and self-confidence. Word of mouth is powerful, and trust and fairness play important roles. With the life circumstances they face, the human aspects are intertwined with the economic as buyers and sellers negotiate while emphasizing both. Transactions can be fluid, with prices and amounts varied (e.g., weighing can be "adjusted" downward if the price bargained is too low) and buyers constantly demanding customization in offerings. As conversations are not private by and large and word can spread very quickly, sellers have to be careful about waiving payment or treating customers differentially.

This world of social relations also cuts both ways and should not be interpreted to suggest anything but an intensely harsh reality. For instance, noncollateral loan repayments are enforced through public humiliation and buyers and sellers are open to abuse and exploitation. Also relevant here are the many different group influences, whether local community groups or larger family or neighborhood or social strata. Group influences can be accentuated in rural areas where villages are geographically dispersed.

Contrasting low-literate consumer behavior in the United States to subsistence consumers in India. In comparing the two consumer groups discussed here, although speculative and cognizant of the wide range of differences within low-literate consumer behavior in the United States, say between urban and rural areas, and similarly in subsistence consumer behavior, a number of differences are noteworthy. In advanced economies like the United States, the marketplace context assumes a certain level of literacy. Large chain stores use technology to compute and present symbolic package and shelf information that assume a certain level of literacy. Thus, low-literate consumers may be isolated from the marketplace and the development of consumer skills may be impeded when compared to the one-to-one interactional subsistence marketplaces described here. In the latter context, generic products are evaluated in face-to-face interactions with prices determined by enquiry and bargaining, and money and

change counted out. Consumers learn skills by being vendors themselves and managing various aspects of a business, rather than through an occupation in a narrowly circumscribed role for a large business as may often be the case in advanced economies. Extreme poverty and the need to get the next meal is in itself a very harsh teacher of consumer lessons. Thus, ironically, the low-literate subsistence consumer may develop some of the functional skills needed to negotiate the one-to-one interactional marketplace, when compared to the low-literate consumer in the United States who has to negotiate a relatively impersonal marketplace that assumes a certain level of literacy.

Implications for Marketing

Each of these streams of research across literacy and resource barriers suggest important implications for marketing. They also represent opportunities for marketers to serve these consumer groups effectively and gain a competitive advantage. Some generic implications that cut across both contexts are noteworthy. First and foremost is the need for literate managers and researchers to understand consumers whom they may not be able to personally relate to in terms of literacy or income. Thus, the most important implication is the need for literate, relatively resource-rich researchers and managers to adopt a mind-set of learning before designing solutions for these contexts. They need to understand consumer behavior in a realm where their own experiences as consumers, although a strength in understanding conventional consumer behavior, may be quite misleading. Thus, the need to employ innovative research methods to understand these consumer groups is critical. Conducting research in these settings requires setting aside preconceptions and employing an open approach that is conducive to learning from low-literate, low-income, or subsistence consumers. Instead of an approach that assumes literate managers and researchers know better and can design solutions for low-literate, low-income consumers or subsistence consumers, an approach of learning from those with experience and expertise in overcoming constraints and then designing solutions for them is likely to be effective.

Low-literate, low-income consumer behavior. A number of implications for marketing relate to addressing the cognitive predilections, emotional elements, decision making and coping discussed earlier (Viswanathan, Rosa, and Harris, 2005; Gau and Viswanathan, 2008). Marketers need to develop an understanding of the challenges faced by low-literate consumers and work to create shopping environments that are user friendly. Examples include price display and sale signs, the use of visuals (e.g., semicircles for half off, dollar bills for amount of savings) rather than abstract symbols such as percentages, and the display of final price, perhaps in a color-coded format, rather than having to rely on computations (Viswanathan, Rosa, and Ruth, 2008). The use of common formats to display information such as unit prices may enable low-literate consumers to inform themselves and begin to use such information to make price-size trade-offs. Pictographic representations such as heart signs or weighing machines as well as graphical representations of nutrient levels, instructions, and other pertinent information are other possible ways to design communications to enable usage by low-literate consumers. Store signs can similarly be designed to depict product categories pictorially and help consumers navigate the store. Familiar pictorial elements are a central means by which low-literate consumers may understand and remember information, therefore, the use of pictorial elements at the package, shelf, and store level are very important as well. Technological aids, such as shopping carts with scanning and computing features, would also be helpful to low-literate consumers as they keep track of their total shopping basket.

Perhaps most importantly, marketers should work to address the challenges faced by low-literate consumers that extend beyond the cognitive to the emotional as well. Training store employees to treat low-literate consumers sensitively and respectfully can go a long way toward building trust and, in turn, loyalty as well, and represents the essence of the marketing approach from which a sustainable competitive advantage can accrue. Perhaps most damaging is the perception among low-literate consumers that they are being cheated, emphasizing the need for employees to be trained and sensitized in these issues, while also distinguishing

between the consequences of low literacy versus other issues such as those faced by literate, English as second language consumers or by literate but poor consumers. Even if low-literate consumers do not complain, the lingering negative emotions may lead them to avoid future visits to such stores. Employees should be trained to interact with different consumers and to explain store policies clearly and respectfully.

At the level of product design, marketers need to learn to use innovative methods to research the needs of low-literate consumers and incorporate insights into the design of the product, including assembly and interface, as well as its packaging and related communications, such as instructions. This requires adopting a different mind-set than one that literate managers and researchers may typically employ, in order to view the usage of products from the perspective of low-literate consumers. A case in point is in designing the interface for computers and cell phones with visual icons, or the packaging of medical products with respect to dosage information.

A noteworthy issue is that many of these implications will ease shopping for all consumers, irrespective of literacy levels. Pictorial aids are easier to use for all consumers, as illustrated by such examples as displays of foods that specific wines can be served with. Product design that incorporates the perspective of low-literate consumers is similarly likely to enhance ease of use for all consumers. With the plethora of information at the store, simplified display of pricing and other information in consistent formats is likely to enhance the shopping experience for all consumers.

Subsistence consumers. Perhaps the biggest implication to marketers is to understand their own limitations in focusing on subsistence consumers owing to their likely lack of personal connection in terms of poverty, literacy, and culture (Viswanathan, Seth, Gau, and Chaturvedi, 2009). Thus, they need to unlearn preconceptions about these consumers and marketplaces and aim to understand their strengths and vulnerabilities as well as marketplace dynamics, such as between buyers and sellers. With an open mind-set and holistic immersion in the context during data collection or a learning and listening phase, and rigorous consideration of what is different about these settings during analysis, marketers in turn, can bring a fresh perspective and new solutions that stem from deep understanding of a radically different context. Thus, complementary strengths and weaknesses that stem from being unfamiliar with subsistence marketplaces can be used to an advantage. Such immersion should aim to understand broader life circumstances and how products and related support would fit in such contexts and improve individual and community welfare. Critical here is the need to avoid exporting solutions from other markets.

Researching subsistence consumers, in turn, involves a number of cognitive, emotional, and administrative considerations (Viswanathan, Gau, and Chaturvedi, 2008). Marketers should gain insights from "experts" in living in subsistence ranging from consumers to vendors, self-help groups, and community-based organizations. Such bottom-up understanding is essential, given the vast differences across different subsistence contexts as a function of geographic location, local language, and cultural differences.

A fundamental challenge for marketers is to identify central needs that subsistence consumers are willing to spend very scarce resources on. Such needs range from the basic to the aspirational, often accompanied by a motivation to pay a small premium for quality – covering food, clothing, and energy, as well as health care, communication and education. In turn, designing solutions requires understanding broader life circumstances as well as specific product-usage situations, often difficult for affluent marketing managers and researchers. How subsistence consumers use or reuse products and use them for multiple purposes is also important to understand. Given the constant demand for customization in a one-to-one interactional marketplace, providing local entrepreneurs with the ability to configure products and prices to different customers is another important consideration (e.g., products with nutritional additives customized to different segments, such as children and the elderly). In designing solutions, the need to involve subsistence consumers and entrepreneurs is also

central, as is the need to address psychological aspects such as engendering trust.

In terms of marketing communications, including informational and educational campaigns, marketers should consider concretizing, localizing, and socializing solutions (Viswanathan, Sridharan, Gau, and Ritchie, 2009). Thus, information in concrete form for low-literate audiences who engage in concrete thinking and pictographic thinking, that also reflect local reality and harness the one-to-one interactional marketplace in terms of community interaction are all likely important elements of marketing communications. This is also the case for product interfaces and packages, with a need to visualize benefits and value propositions. In contexts of multifaceted deprivation, marketers should also consider product-relevant support including educational programs. Another important issue is the need to communicate the value proposition to highlight costs (e.g., including time and effort) and benefits (including hidden benefits such as nutrition to avoid illnesses) (Sridharan and Viswanathan, 2008). Such costs and benefits and the value proposition need to be communicated to low-literate audiences in ways that can be visualized. Communications should also be designed from the bottom up to harness the social networks discussed here, through partnerships with local entrepreneurs and customers. Similarly, distribution should harness social networks, such as through local retailers and community-based organizations and allow for customization at the point of purchase through such means as offering credit.

To implement their plans, marketers need to work with a diverse set of organizations who have been functioning in subsistence marketplaces, in a number of sectors of society, such as local governments and community-based organizations. Lacking the institutionsin affluent markets that enable impersonal exchanges, relationships with diverse groups and organizations are critical in providing access and effectively implementing plans.

Marketers should also understand how product performance and related company activities work toward improving individual and community welfare (Viswanathan *et al.*, 2009). For instance, subsistence consumers seek products that will better their life circumstances, ranging from nutritious food, to education for their children and communication devices. When products lack these characteristics, subsistence consumers may opt to make or forgo the products rather than spend scarce resources. Therefore, quality and value are extremely important. Moreover, product support through educational programs and other means may be critical in settings where deprivation is multifaceted. In their quest to survive, the human aspects and larger life circumstances are intertwined with economic transactions – in contrast to impersonal transactions in affluent societies with relative certainty regarding infrastructure and institutional support. Thus, issues of trust and fairness are very important and marketers should work to address them. With the one-to-one interactional marketplace described earlier characterized by strong word of mouth, a reputation for trustworthiness and fairness can go a long way in developing a sustainable competitive advantage. Focusing on individual and community welfare may also be helpful in working with diverse organizations to reach and serve markets.

In conclusion, the need to view subsistence consumers as preexisting marketplaces whose dynamics need to be understood by marketers should be emphasized. Rather than view these subsistence marketplaces as parallel markets to sell to, they should be viewed as marketplaces to learn from, to then design solutions for, that may well be transferable to all contexts. As a case in point, creating products for conditions of very scarce recourses and lack of infrastructures, such as, say, distributed energy generators (solar applications), in turn has implications for all contexts in confronting the challenges of the twenty-first century.

In summary, a top-down approach to different elements of marketing assumes certain institutional infrastructure, often lacking in subsistence marketplaces. For instance, marketing research assumes the ability to sample, marketing communications assume mass media, and distribution assumes related infrastructure. With the lack of such infrastructure, subsistence marketplaces rely on social networks to provide some measure of certainty in an extremely uncertain world. Rather than adopt a mind-set of scaling by

using top-down approaches, marketers should employ a bottom-up approach to addressing each element of marketing in radically different and diverse subsistence contexts, and aggregate insights from the bottom up.

In conclusion, this article explored consumer behavior across literacy and resource barriers, in sharp contrast to conventional consumer behavior. Both for low-literate, low-income consumers in the United States, and for subsistence consumers in South India, traditional assumptions about consumer behavior do not hold. Rather, unique cognitive, emotional, and social characteristics that characterize such consumer behavior present marketers with challenges as well as opportunities to serve such consumers and develop a sustainable competitive advantage while also improving individual and community welfare.

Bibliography

Adkins, N. and Ozanne, J. (2005) The low literate consumer. *Journal of Consumer Research*, **32** (1), 93–105.

Alwitt, L. (1995) *The Low-Income Consumer: Adjusting the Balance of Exchange*, Sage, Thousand Oaks.

Arnould, E.J. and Mohr, J.J. (2005) Dynamic transformations for base-of-the-pyramid market clusters. *Journal of the Academy of Marketing Science*, **33** (3), 254–274.

Bettman, J.R. (1979) *An Information Processing Theory of Consumer Choice (Duke University)*, Addison-Wesley Publishing Co., Reading.

Gau, R. and Viswanathan, M. (2008) The retail shopping experience for low-literate consumers. *Journal of Research for Consumers*, (15), Consumer Empowerment Special Issue.

Hill, R.P. (1991) Homeless women, special possessions, and the meaning of 'home': an ethnographic case study. *Journal of Consumer Research*, **18** (3), 298–310.

Kirsch, I.S. and Guthrie, J.T. (1997) The concept and measurement of functional literacy. *Reading Research Quarterly*, **13** (4), 485–507.

Kutner, M., Greenberg, E., and Baer, J. (2005) *A First Look at the Literacy of America's Adults in the 21st Century*, National Center for Education Statistics, Department of Education, Washington, DC.

Luria, A.R. (1976) *Cognitive Development: Its Cultural and Social Foundations*, Harvard University Press, Cambridge.

Prahalad, C.K. (2005) *The Fortune at the Bottom of the Pyramid*, Wharton School Publishing, University of Pennsylvania, Philadelphia.

Sridharan, S. and Viswanathan, M. (2008) Marketing in subsistence marketplaces: consumption and entrepreneurship in a South Indian context. *Journal of Consumer Marketing*, Special Issue on Base of the Pyramid Research, **25** (7), 455–462.

Viswanathan, M. (2007) Understanding product and market interactions in subsistence marketplaces: a study in South India, in *Product and Market Development for Subsistence Marketplaces: Consumption and Entrepreneurship Beyond Literacy and Resource Barriers*, Advances in International Management Series (eds J. Rosa, M. Viswanathan, J. Cheng, and M. Hitt), Elsevier, Oxford, pp. 21–57.

Viswanathan, M., Gajendiran, S., and Venkatesan, R. (2008) *Enabling Consumer and Entrepreneurial Literacy in Subsistence Marketplaces*, Springer, Dordrecht.

Viswanathan, M., Gau, R., and Chaturvedi, A. (2008) Research methods for subsistence marketplaces, in *Sustainability Challenges and Solutions at the Base-of-the-Pyramid: Business, Technology and the Poor* (eds P. Khandachar and M. Halme), Greenleaf Publishing, Sheffield.

Viswanathan, M., Rosa, J.A., and Harris, J. (2005) Decision-making and coping by functionally illiterate consumers and some implications for marketing management. *Journal of Marketing*, **69** (1), 15–31.

Viswanathan, M., Rosa, J., and Ruth, J. (2008) *Emerging Lessons – For multinational companies, understanding the needs of poorer consumers can be profitable and socially responsible*, Wall Street Journal/MIT Sloan Online, October.

Viswanathan, M., Rosa, J.A., and Ruth, J., (2010) Exchanges in marketing systems: the case of subsistence consumer merchants in Chennai, India. *Journal of Marketing*, **74**, 1–18.

Viswanathan, M., Seth, A., Gau, R., and Chaturvedi, A. (2009) Internalizing social good into business processes in subsistence marketplaces: the sustainable market orientation. *Journal of Macromarketing*, **29**, 406–425.

Viswanathan, M., Sridharan, S., Gau, R., and Ritchie, R. (2009) Designing marketplace literacy education in resource-constrained contexts: implications for public policy and marketing. *Journal of Public Policy & Marketing*, **28** (1), 85–94.

Viswanathan, M., Xia, L., Torelli, C., and Gau, R. (2009) Literacy and consumer memory. *Journal of Consumer Psychology*, **19**, 389–402.

Zaichkowsky, J.L. (1985) Measuring the involvement construct. *Journal of Consumer Research*, **12**, 341–352.

emotion

Julie A. Ruth

Overview and Definition of Emotions in Consumer Behavior

Emotions are psychological states of readiness that encompass thoughts, subjective feelings, physiological changes, expressive behaviors, and action tendencies (Bagozzi, Gopinath, and Nyer, 1999). The specific character of the subjective feeling – for example, love, pride, excitement, anxiety, fear, guilt, or anger – is associated with systematic patterns of appraisals the consumer makes about the event in the light of implications for well-being. In addition to appraisals, emotions have a felt, experiential component that is accompanied by physiological processes such as heightened arousal. Emotions are often communicated physically through gestures or facial expressions, allowing emotions to spread to others, and may result in the consumer taking specific actions to adapt to the situation at hand.

Consistent with approaches in psychology, consumer research distinguishes between emotions and related constructs such as affect and mood. Emotions typically have a target, such as the consumer being exposed to a TV ad that elicits guilt or a consumer experiencing anger in light of a provider's inept service delivery. In contrast, affect is conceptualized as a feeling state that is largely undifferentiated beyond its positive or negative valence, or as an umbrella term that encompasses a set of more specific feeling phenomena such as emotions, moods, and/or attitudes. Moods are conceptualized as lower in intensity than emotions and are less differentiated due to their lack of a specific target and lack of correspondence with specific appraisals and action tendencies (Bagozzi, Gopinath, and Nyer, 1999; Cohen, Pham, and Andrade, 2008).

Emotions can be experienced in anticipation of, during, or after consumption episodes such as fear in response to an ad identifying risks to family safety, delight and regret upon making a good or poor decision respectively, or anger when receiving poor or prejudiced customer service. Consumers can also experience a mix of emotions such as fear, joy, regret, and excitement in conjunction with skydiving or viewing horror movies. Using frameworks of emotion developed in psychology as a foundation, Richins (1997) developed a set of descriptors reflecting the emotions consumers may experience when anticipating, acquiring, or possessing and using products. The Consumption Emotion Set (CES) includes multiple scale items associated with 16 emotions: anger, discontent, worry, sadness, fear, shame, envy, loneliness, romantic love, love, peacefulness, contentment, optimism, joy, excitement, and surprise. Richins also shows that consumers associate different emotions with different types of possessions. For example, sentimental objects such as heirlooms and mementos are associated with love, and recreational products such as stereo equipment are associated with positive emotions including excitement. In contrast, vehicles are associated with a combination of positive and negative emotions such as joy, pride, anger, excitement, worry, and guilt.

The next section describes the ways in which appraisals, coping, and goal striving are related to consumers' experience of emotions.

Consumers' Experience of Emotions

Appraisals. Emotions typically arise in conjunction with appraisals consumers make in situations of personal significance. Appraisals are evaluative judgments and interpretations of the meaning and importance of elements of the situation at hand. Appraisals may be conscious or may occur automatically upon perception. Johnson and Stewart (2005) synthesize appraisal theories drawn from psychology and suggest that six appraisals differentiate and inform the experience of discrete consumer emotions: (i) direction of goal congruence (i.e., positive or negative); (ii) agency, or the extent to which the locus of responsibility and control resides with the consumer, some other entity, or beyond anyone's responsibility and control; (iii) certainty; and (iv) normative or moral compatibility. In addition, appraisals of (v) goal importance and (vi) degree of goal congruence are linked to the intensity of emotional experience.

Different emotions are characterized by systematic patterns of appraisals. For example, as shown in research on gift exchange, gift recipients experience love in pleasant, goal-congruent

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

gift-exchange situations in which certainty and other-agency are relatively high and self-agency is low. In contrast, recipients experience anger in unpleasant, goal-incongruent exchange circumstances in which certainty and other-agency are high and self-agency is low. It is also important to note that, although there is evidence that appraisals can be causal in their influence on emotions, other mechanisms can elicit emotions. For example, noncognitive methods such as bodily feedback or unconscious priming have also been shown to elicit specific emotions.

Coping. Emotion regulation is aimed at coping with the implications of a given situation for the consumer and his or her goals and well-being (*see* MOTIVATION AND GOALS; CONSUMER WELL-BEING). Coping is the means by which an individual identifies and assesses the adaptive potential and significance of various actions. Of critical importance is the assessment the consumer makes in light of the actual situation compared to the desired one. Two appraisals are particularly important at this stage of emotion formation: goal relevance and goal congruence, or the extent to which the consumer perceives a personal stake in the situation at hand and the degree to which the event facilitates or inhibits this stake.

A prominent view suggests that two broad categories of coping strategies are available. Problem-focused coping aims at maintaining the original goal and taking action to attain the goal or maintain its achievement. Emotion-focused coping, on the other hand, involves an attempt to reframe the situation by shifting appraisal(s) and/or revising the goal. Such revisions in appraisal or goals shift the emotional reaction in light of the situation. For example, the angry consumer who lodges a complaint in order to resolve a negative situation may be activating problem-focused coping. In contrast, a consumer may be able to shift emotions from anger to disappointment through use of an emotion-focused coping strategy if he/she reframes the situation by strengthening beliefs that the service provider was not in control of a negative outcome.

Recent research shows that consumers frequently rely on both emotion-focused and problem-focused coping strategies and may do so even within the same consumption episode. Further, some coping strategies may satisfy both emotion- and problem-focused coping goals. Duhachek (2005) reviews the psychological literature on coping and provides evidence for three broad categories of coping strategies used by consumers: (i) cognitive or behavioral (e.g., engaging in rational thinking about the situation, trying to make the best of the situation through positive thinking, or taking action to resolve the situation); (ii) expressive support seeking (e.g., engaging in emotional venting to others, seeking instrumental assistance from others, or seeking emotional support from others); and (iii) avoidance coping (e.g., avoiding the situation or denying it). Duhachek also shows that coping strategies are dependent on the specific emotion the consumer experiences and the consumer's beliefs about coping efficacy. For example, high self-efficacy consumers show a tendency to enact expressive support-seeking coping strategies when experiencing fear in contrast to active coping strategies when experiencing anger.

Goal striving and action tendencies. Closely related to coping is action tendency, which is a readiness to engage in or disengage from interaction with an object to bring the situation in line with goals. Anticipated emotions serve a role in goal striving through prompting intentions, plans, and decisions of how to allocate resources in order to move toward goal attainment (*see* MOTIVATION AND GOALS; CONSUMER INTENTIONS). Specific emotions also provide consumers with valuable feedback about their progress toward goal attainment. Happiness, for example, is perceived by consumers as a signal that a goal has been achieved, whereas sadness is a signal that goal achievement has not been attained.

Much as emotions arise in response to patterns of appraisals, there is evidence that patterns of action readiness also correspond with distinct emotions. Although positive (negative) consumption emotions may stimulate positive (negative) behavioral intentions, their shared valence is not the only factor that shapes behaviors. The nature of the consumer's goals also shapes action tendencies. Pride, for example, is associated with goal attainment, but it can occur in the light of achieving either

promotion or prevention self-regulatory goals. Research shows that consumers who experience pride in conjunction with prevention goals (e.g., striving to avoid paying extra money for a purchase) are less likely to repurchase than those with promotion goals (e.g., striving to obtain a discount as means to gain money), whereas there are no differences in repurchase behavior among consumers who experience low levels of pride in conjunction with either prevention or promotion goals.

The Role of Emotion in Four Domains of Consumer Activity

Emotions play important roles in many domains of consumer activity. Four key domains are consumer response to advertising, consumer decision making, purchase and service situations, and possession and usage.

Emotions and advertising persuasion. The majority of research on the role of emotion in advertising has focused on two aspects of persuasion: (i) the nature and effects of emotions evoked in response to ads; and (ii) how consumers process and evaluate ads that appeal on the basis of emotion (*see* ADVERTISING MESSAGE APPEALS; PERSUASION; ATTITUDES; ADVERTISING EFFECTIVENESS).

Ad-evoked feelings have been shown to systematically shape attitudes the consumer develops and holds toward the ad and brand. For example, Edell and Burke (1987) present a scale measuring emotional responses toward ads and find three factors underlying these responses: upbeat feelings such as delight, happiness, and pride; negative feelings such as anger and irritation; and warm feelings such as hope and having been moved by the ad. These positive and negative emotional responses to ads can co-occur and, along with cognitive responses, influence the formation of attitudes toward the ad and indirectly influence brand attitudes. Emotional responses to ads have been found to be more important determinants of attitude toward the ad than thoughts under low involvement conditions, whereas both cognitive and emotional responses are important under high involvement conditions.

Turning to consumer response to emotional appeals, these ads affect persuasion through eliciting a sufficiently high level of emotion that attracts attention and prompts consumer processing of the message, which in turn contributes to behavioral intentions. For example, research has shown that negative emotional appeals can prompt positive outcomes, as when a negative emotional ad highlights the needs of abused children. These types of prosocial appeals elicit consumer feelings of sadness, anger, fear, and empathy. Stronger felt negative emotions tend to prompt greater feelings of empathy, which in turn facilitate intentions to help.

An inverted U-shaped pattern of effects of emotional appeals on persuasion is sometimes observed. In such instances, attitudes toward the ad and brand are relatively favorable under moderate levels of the elicited emotion and less favorable under either low or high levels of emotional response. For example, research shows that fear appeals that elicit low levels of the emotion are low in persuasion because they prompt insufficient elaboration of the threat or harm referred to in the ad. Likewise, appeals that elicit high levels of fear are also low in persuasion because they prompt an overly high level elaboration of the threat or harm that interferes with processing the recommendations for action to alleviate the fear. In contrast, appeals calibrated to elicit a moderate level of the emotion appear to strike an effective persuasive balance because the consumer is able to both elaborate on the threat and process the recommended action. Interventions that enhance the likelihood of elaboration, such as self-referencing, can increase persuasiveness of low-fear appeals, and elaboration-reducing interventions can likewise increase the persuasiveness of high-fear appeals (Keller and Block, 1995). Some research has also observed a similar pattern for appeals intended to elicit positive emotions, such as humorous appeals. Low levels of humor can be weak in attracting consumer attention, and very high levels of humor can distract the consumer from processing the marketing message. Moderate levels of positive emotion elicited through humor appears to strike an effective balance in attracting attention and prompting processing while leading to positive attitude formation.

Appraisal-related dimensions of emotional ads also influence persuasion. Whether an ad refers to high or low consumer agency, for example, shapes male consumers' attitudes toward the ad. Males form less favorable attitudes toward ads depicting a low self-agency emotion, such as peacefulness or tenderness, when viewing with another male. This result is attributed to the incongruence between the low-agency emotion and male stereotypes, which are made salient by the viewing circumstances. In contrast, ad attitudes are more favorable when males are exposed to either a low-self agency ad with a female viewing partner or when exposed under any viewing condition to an ad depicting a stereotype-congruent, high-agency emotion such as excitement or joy. Females do not display differential responses to high-versus low-agency emotional ads.

Members of different cultural groups may also process and evaluate emotion-based ads in systematically different ways. Consumers from collectivist cultures rather than individualistic ones have more favorable attitudes toward ego-focused (e.g., pride, happiness) versus other-focused (e.g., empathy, peacefulness) emotional appeals and brands. Consumers' motivation to process explains this somewhat counterintuitive pattern of results. That is, ads that are incongruent with cultural values are relatively novel and thus attention getting, which yields greater liking; ads that are easier to process because the appeals are consistent with cultural values are less novel and hence less attractive.

Some ads contain a mix of positive and negative emotional appeals, such as those combining happiness and sadness. Williams and Aaker (2002) find evidence that mixed appeals yield less favorable attitudes among consumers with a lower propensity to accept duality compared to those with a higher propensity, such as Anglo-Americans and Asian-Americans, respectively. Liking and recall of emotional ads tends to be higher among older compared to younger consumers. Further, ads focusing on avoiding negative emotions are liked and recalled more by older consumers and also younger consumers who hold a limited compared to longer time horizon.

Emotions and decision making. Emotions have been found to exert a variety of influences in consumer decision making through the signals they provide with respect to the consumer's movement toward goals and managing the situation at hand (*see* CONSUMER DECISION MAKING). Whereas considerable research examines ambient affect as a context effect (see Cohen, Pham, and Andrade, 2008), other research investigates a more central role of emotions in decision-making processes. According to an "affect-as-information" model, consumers may evaluate products by holding a target representation in mind and asking, "How do I feel about it?" From this theoretical perspective, consumers' emotions are considered to be valid sources of information.

When holding the target experience in mind, such as a consumer considering, "Should I go shopping this afternoon?", a target that elicits positive feelings leads to more favorable evaluations than a target that elicits negative emotions. The extent to which consumers rely on such feelings in decision making depends on (i) the heuristic value of feelings, (ii) their representativeness, and (iii) their perceived relevance (Pham, 1998). Specifically, feelings are more likely to be used when feelings have value as a type of heuristic, such as situations where few other sources of information are available or when feelings can be used to simplify the decision process. For example, the consumer may elect to eliminate from a choice set all alternatives that are not associated with positive feelings. Whether consumers rely on such feelings also depends on the extent to which feelings are believed to be representative of the target. Consumers are less likely to rely on them if they believe the feelings have not been elicited by the target itself, as when happiness is associated with the weather rather than the target consumption experience. Finally, consumers tend to rely on feelings when they are regarded as more relevant to the target, such as situations where consumers have experiential motives compared to instrumental ones. Consumers who tend to process in a sensory or visual manner also display a tendency to act in accordance with the affect-as-information framework.

While research has largely observed a carry-over effect of emotional valence on judgment and

choice, more recent studies argue that specific emotions give rise to specific cognitive and motivational processes that are systematically related to patterns of effects of emotions on decision making. Specifically, consumer emotions of the same valence, such as anger and sadness, can have different effects on judgment and choice because key appraisals and core experiential themes associated with them differ. Likewise, emotions of different valences may even have similar effects on judgments if they share other key appraisals and core themes.

Still, many choice decisions are in and of themselves difficult for consumers and elicit negative emotions. In their model of choice trade-off difficulty, Luce, Bettman, and Payne (2001) suggest that consumers appraise choice situations in light of goals and emotional content. They show that negative task-related emotions arising in situations involving difficult emotional trade-offs lead to various forms of coping that are either directed toward the problem or the emotion. If processing resources are limited, emotional reactions that are evoked spontaneously by an alternative in the choice set tend to have a greater impact on decision making and choice than cognitive reactions. In these constrained processing conditions, consumers tend to choose alternatives that are superior on an emotion-related rather than a cognitive dimension.

In contrast to trade-off difficulty, in other instances consumers may doubt whether what they yearn for is possible and thus find that their feelings of hope are threatened (de Mello, MacInnis, and Stewart, 2007). Consumers experience hope as a positive-valenced emotion that is elicited in situations that are uncertain but hold the possibility of achieving a goal-congruent outcome. If consumers lose confidence in a hoped-for outcome, they tend to engage in motivated reasoning when processing information about a product touted to enable goal achievement. That is, they tend to produce more self-serving product judgments, seek out product-supportive information sources, make biased assessments of information credibility, exhibit weaker distinctions between high- and low-credibility product information, and place less weight on negative information.

Kidwell, Hardesty, and Childers (2008) recently developed the Consumer Emotional Intelligence Scale (CEIS) to measure individual differences in consumers' ability to skillfully use emotion-based information in information processing and decision making (*see* CONSUMER INFORMATION PROCESSING; CONSUMER DECISION MAKING; EMOTIONAL INTELLIGENCE). The scale consists of items measuring four underlying dimensions of consumer emotional intelligence: *perceiving* emotions accurately; *facilitating*, or the ability to access and use emotions in mental processes; *understanding* emotions and their meaning; and *managing* emotions in the process of achieving desired outcomes. Consumers with higher compared to lower levels of emotional ability and confidence in this ability tend to make higher compared to lower quality choices.

Emotions arising in purchase and service settings. **Psychological aspects of emotions in purchase and service settings.** Emotions that arise in product purchase or service delivery shape the consumer's assessments of satisfaction, complaint behavior, and word-of-mouth communications (*see* CONSUMER BEHAVIOR AND SERVICES MARKETING; CUSTOMER SATISFACTION). Research has largely found that the experience of positive emotions increases satisfaction while the experience of negative emotions decreases satisfaction, over and above the influence of expectations, product performance, and disconfirmation processes. Products that meet or exceed consumers' hedonic wants and fulfill promotion goals also tend to increase delight, which in turn contributes toward more favorable word-of-mouth and repurchase intentions.

Negative emotions arising from purchase have been the subject of numerous studies. Regret, for example, is experienced when the consumer makes the judgment that a foregone alternative performs better than the chosen alternative (Tsiros and Mittal, 2000). Consumer regret and satisfaction have been shown to have different antecedents, moderators, and consequences. Regret tends to be experienced when the consumer recognizes a better but forgone option and engages in counterfactual thinking in conjunction with a chosen outcome that is

negative and irreversible. Whereas regret and satisfaction both directly influence repurchase intentions, the influence of regret on complaint intentions is mediated by satisfaction. Although research has generally found that consumers experience more regret when deciding to make a change from the status quo compared to maintaining it, feelings of regret can be mitigated if the consumer reflects and concludes that the decision was appropriate under the circumstances.

There is evidence of systematic relations between appraisals, emotions, and the coping strategies that consumers employ in negative purchase situations. Yi and Baumgartner (2004) find that two dimensions explain the appraisal, emotion, and coping patterns of consumers experiencing anger, disappointment, regret, or worry in purchase: (i) the extent to which the consumer believes the problem can or cannot be managed; and (ii) the extent to which the consumer can manage or not manage the emotion itself. In anger experiences, where responsibility and control are attributed to another party and the situation is deemed changeable, a high problem-focus/low emotion-focus coping strategy of confrontation is likely to be used. In disappointing purchase situations associated with unmet expectations and a situation that is beyond the control of the consumer or other party, low problem-focus/low emotion-focus coping strategies such as disengagement are likely to be enacted. When consumers blame themselves and experience regret in purchase, they tend to engage in acceptance and positive reinterpretation coping strategies, which are low in problem-focus and high in emotion-focus. Worry, on the other hand, arises in consumption situations where the future is uncertain, and so a variety of coping strategies are enacted, depending on perceptions of control over the situation.

Numerous studies have also examined negative emotion experienced in service failures (*see* CONSUMER BEHAVIOR AND SERVICES MARKETING). Consumers make assessments about the characteristics of the failure, which are in turn associated with different emotions such as anger or regret. When consumers appraise a goal-incongruent situation as having been caused by the service provider, consistent with its characteristic appraisal and action tendency patterns, the consumer is more likely to experience anger than regret and exhibit more retaliatory behaviors. Whereas marketers' recovery efforts that are designed to reduce consumer anger are associated with a reduction in consumers' retaliatory behavior, recovery efforts that aim to shift blame from the service provider to the consumer typically result in increased consumer anger and retaliatory behavior. Moreover, consumers' negative emotions such as anger or exasperation are highly predictive of their negative word-of-mouth behavior. Consumers' positive emotions, in contrast, are not related to their negative word-of-mouth behavior in these situations.

Social aspects of emotions in purchase and service settings. Many purchase and service contexts involve the presence of other people. How does that social situation affect consumer emotions? How is the consumer affected by another person and that individual's emotional state? Some research argues that emotional contagion occurs when one person "catches" the emotion being experienced and expressed by another person, so that the emotion of the receiver converges with that of the sender. Recent findings suggest that the role played by an emotion sender – a fellow consumer versus an employee – and the authenticity of the sender's emotional display influence how the emotional message is received by consumers. For example, Howard and Gengler (2001) find that emotional contagion occurs when happiness transfers from senders to receiving consumers who have positive relational bonds with senders. The facial expression of senders must be visible in order for receiving consumers to mimic the smiling and experience the happy emotion, which in turn has a positive impact on receivers' attitudes toward products that are present in the situation. Research on employee–customer interactions shows that employees' emotional displays can also trigger changes in customers' emotions but that, in addition to the bond established between employees and customers, the correspondence in employee/customer emotions may depend on the authenticity of employees' emotional display rather than the mere act of smiling itself.

Consumer emotions are also affected by brief interactions with unacquainted others or even a noninteractive "mere" presence of others in a shopping environment. Dahl, Honea, and Manchanda (2005) show that consumers who interact briefly with a salesperson compared to those who do not are more likely to experience guilt if they do not ultimately make a product purchase. Consumers are also more likely to experience embarrassment in purchasing an embarrassing product in the presence of another. Several other factors affect the degree to which consumers experience embarrassment in purchasing embarrassing products: the frequency of the potentially embarrassing situation's occurrence, the severity of the social threat, and the consumer's public self-consciousness. Likewise, the presence or absence of others can affect the experience of anger in service failure situations. Consistent with research on the salience of felt ethnicity in numerical minority social conditions, Baker, Meyer, and Johnson (2008) find that service failures elicit higher levels of anger among black consumers when no other black customers are present compared to situations where several other black consumers are present, or compared to white consumers in either condition.

Emotions arising in possession and usage. Studies examining various aspects of possession and usage have found emotions to be central to these experiences in two noteworthy ways. First, mixed emotions are common because consumption experiences almost always involve multiple and sometimes conflicting attributes and goals. Second, possession and usage contexts are often social in nature, and so interpersonal and social factors may be more salient in the experience and effects of emotions in these types of consumption settings.

The recent work on indulgent consumption acknowledges the importance of positive and negative emotions, showing that their effects are dependent on the specific combination of mixed emotions that consumers experience. With indulgent consumption, both prudent and impulsive consumers experience a mix of positive and negative emotions such as excitement and frustration. Yet prudent, but not impulsive, consumers also tend to experience negative self-conscious emotions such as guilt and shame. As time elapses after indulging, impulsive consumers are more likely to feel residual effects of their positive emotions as their negative feelings dissipate, whereas negative emotions including self-conscious feelings tend to persist for prudent consumers while their positive emotions dissipate. As a result, prudent compared to impulsive consumers are less likely to indulge when given a subsequent opportunity to do so (*see* IMPULSIVE AND COMPULSIVE BUYING).

The nature of emotions experienced in consumption also affects memory, where mixed consumption experiences involving both positive and negative emotions are more difficult to recall accurately than unipolar ones. In addition, over time consumers remember mixed emotion experiences as less mixed, a finding that highlights a type of memory decay that does not occur with unipolar emotion experiences. The felt conflict of mixed emotions appears to explain this decay. The memory bias is also more evident among Anglo-American compared to Asian-American consumers, consistent with other research regarding cultural differences regarding acceptance of duality.

At first glance, consumers' pursuit of consumption experiences laden with negative emotions, such as horror movies, would seem to be counterintuitive. Research shows that viewing horror movies elicits both negative and positive emotions in consumers predisposed to viewing horror movies. In contrast, only negative emotions are activated among consumers who typically avoid these types of consumption activities. Although similarly high levels of negative emotions are felt by both types of individuals, consumers with a protective orientation that provides psychological distance or detachment are able to experience positive emotions as well (see Cohen, Pham, and Andrade, 2008).

Turning to social aspects of consumption, some research has examined the correspondence between the emotions consumers experience and the relationships they hold with others including family, friends, and acquaintances. For example, gift exchange oftentimes elicits a mix of emotions, such as joy and love in being acknowledged as a gift recipient but

also the anxiety a recipient may experience in being the center of attention and obligated to conform to dramatic scripts associated with rituals like baby showers or holiday observances. As a result, valence alone does not explain the relationship impact of gift exchange. For example, research shows that gift recipients who perceive that the relationship with the giver is strengthened typically feel a mix of positive and negative emotions such as joy in the relationship heading in a desired direction but fear of the unknown that lies ahead in the relationship. Recipients who believe already-strong relationships are affirmed typically experience positive emotions in gift-exchange processes (Ruth, Otnes, and Brunel, 1999).

With a focus on participation in rituals, Otnes, Lowrey, and Shrum (1997) explore the nature of consumer ambivalence, defined as the simultaneous or sequential experience of multiple emotional states that are shaped by the consumer's interaction with social or cultural phenomena in market-oriented contexts and that influence prepurchase, purchase, or postpurchase attitudes and behavior. Four antecedents of consumer ambivalence emerge in their study of consumer participation in wedding planning: a gap between expectations and actual experience, overload in decision-making processes, role conflict with others, and conflicts among customs and/or values. These antecedents are associated with specific coping strategies that consumers use to manage their ambivalent feelings such as striving to assertively resolve gaps between expectations and reality regarding products or retailer services, or compromising to deal with role conflicts.

Looking Forward

Much of the extant research in our field has adopted a psychological perspective, concerned with the individual consumer's experience of emotion and its effect on other individual-level phenomena such as the consumer's attitudes, choices, assessments of satisfaction, and goals (*see* ATTITUDES; CONSUMER DECISION MAKING; CUSTOMER SATISFACTION; MOTIVATION AND GOALS). Yet, it is important to note that consumer behavior frequently occurs within interactions with service providers, friends, family, and even strangers in servicescapes and shopping environments. Building on extant research with a psychological foundation, recent studies of the social aspects (*see* SOCIAL INFLUENCE) of emotional experience and expression provide a strong platform for expanding research on how consumer emotions are elicited and shared between and among consumers and marketing agents such as service providers and sales personnel. Such initiatives could address how such emotions affect interpersonal phenomena, for instance, assessments of interpersonal relationship quality, trust, and commitment.

Such initiatives could also contribute to our understanding of how consumer emotions contribute to and shape the meanings and emotional attachments consumers have with brands and BRAND COMMUNITY, the ways in which consumers experience and express nostalgia, and other emotion-based aspects of possession and usage. In addition, recent conceptualizations emphasizing the perspective that emotions are informative and help consumers to function and adapt to changing situations provide promising directions for new knowledge insights on emotions in consumer behavior.

Bibliography

Bagozzi, R.P., Gopinath, M. and Nyer, P.U. (1999) The role of emotions in marketing. *Journal of the Academy of Marketing Science*, **27** (2), 184–206.

Baker, T.L., Meyer, T. and Johnson, J.D. (2008) Individual differences in perceptions of service failure and recovery: the role of race and discriminatory bias. *Journal of the Academy of Marketing Science*, **36** (4), 552–564.

Cohen, J.B., Pham, M.T. and Andrade, E.B. (2008) The nature and role of affect in consumer behavior, in *Handbook of Consumer Psychology* (eds C.P. Haugtvedt, P.M. Herr and F.R. Kardes), Lawrence Erlbaum, New York, pp. 297–348.

Dahl, D.W., Honea, H. and Manchanda, R.V. (2005) Three Rs of interpersonal consumer guilt: relationship, reciprocity, reparation. *Journal of Consumer Psychology*, **15** (4), 307–315.

de Mello, G.E., MacInnis, D.J. and Stewart, D.W. (2007) Threats to hope: effects on reasoning about product information. *Journal of Consumer Research*, **34** (2), 153–161.

Duhachek, A. (2005) Coping: a multidimensional, hierarchical framework of responses to stressful consumption episodes. *Journal of Consumer Research*, **32** (1), 41–53.

Edell, J.A. and Burke, M.C. (1987) The power of feelings in understanding advertising effects. *Journal of Consumer Research*, **14** (3), 421–433.

Howard, D.J. and Gengler, C. (2001) Emotional contagion effects on product attitudes. *Journal of Consumer Research*, **28** (2), 189–201.

Johnson, A.R. and Stewart, D.W. (2005) A reappraisal of the role of emotion in consumer behavior: traditional and contemporary approaches, in *Review of Marketing Research* (ed. N.K. Malhotra), M. E. Sharpe, Armonk, NY, pp. 1–33.

Keller, P.A. and Block, L.G. (1995) Increasing the persuasiveness of fear appeals: the effect of arousal and elaboration. *Journal of Consumer Research*, **22** (4), 448–459.

Kidwell, B., Hardesty, D.M. and Childers, T.L. (2008) Consumer emotional intelligence: conceptualization, measurement, and prediction of consumer decision making. *Journal of Consumer Research*, **35** (1), 154–166.

Luce, M.F., Bettman, J.R. and Payne, J.W. (2001) *Emotional Decisions: Tradeoff Difficulty and Coping in Consumer Choice*, University of Chicago Press, Chicago.

Otnes, C.C., Lowrey, T.M. and Shrum, L.J. (1997) Toward an understanding of consumer ambivalence. *Journal of Consumer Research*, **24** (1), 80–93.

Pham, M.T. (1998) Representativeness, relevance, and the use of feelings in decision making. *Journal of Consumer Research*, **25** (2), 144–159.

Richins, M.L. (1997) Measuring emotions in the consumption experience. *Journal of Consumer Research*, **24** (2), 127–146.

Ruth, J.A., Otnes, C.C. and Brunel, F.F. (1999) Gift receipt and the reformulation of interpersonal relationships. *Journal of Consumer Research*, **25** (4), 385–402.

Tsiros, M. and Mittal, V. (2000) Regret: a model of its antecedents and consequences in consumer decision making. *Journal of Consumer Research*, **26** (4), 401–417.

Williams, P. and Aaker, J.L. (2002) Can mixed emotions peacefully coexist? *Journal of Consumer Research*, **28** (4), 636–649.

Yi, S. and Baumgartner, H. (2004) Coping with negative emotions in purchase-related situations. *Journal of Consumer Psychology*, **14** (3), 303–317.

consumer behavior and services marketing

Dawn Iacobucci

Introduction

Let us take a simple consumer behavior process and see how it is studied and modified by services marketing scholars. In particular, let us begin with the basic, vanilla-flavored consumption model, in which a consumer (i) identifies some desire, (ii) searches for means of satisfying that desire, (iii) makes a choice and a purchase, and (iv) afterward reflects upon that purchase. This article elaborates on the stages of this process with an eye toward identifying the special concerns of the services marketing researcher.

What is Service Marketing?

Presumably, the reader of this volume will walk away with a strong and clear understanding and appreciation of the explication of the consumption process. What is new to this article is the twist on services. Hence, a starting point is to clarify what is meant by services marketing. While the phrase *services marketing* is rather general, it tends to mean one of two things: a researcher is studying either customer service or some consumption phenomenon in a service industry.

The breadth of services marketing is part of what makes teaching services fun because the cases and applications run the gamut from customer service in, say, an automobile dealership, where the focus of the purchase is a good, or a call center, where the focus is often on service recovery, to fairly routine and familiar services such as hotels and restaurants, to the more sophisticated services such as in the health care or legal professions.

The range of applicability can be explained in part from the philosophy that just about every purchase contains some element of service. In the early days, services marketing scholars made a point of distinguishing services (e.g., dry cleaning) from goods (e.g., shampoo) to highlight the conceptual differences worthy of study. Contemporary thinking is that very few things exist that have not been processed in some manner (e.g., perhaps coal or wheat), and that the value-added processing is itself service. This argument of the pervasiveness of services seems to be getting more difficult to challenge as so much of the value-added element is information-based, both in B2C and B2B markets.

What is Different about Services?

There are several dimensions along which services have been distinguished from goods, or along which the service element of a purchase can be distinguished from whatever goods might be transacted in the purchase. The acronym "SHIP" is helpful to remember the primary dimensions: the "S" stands for "simultaneity of production and consumption," the "H" for "heterogeneity," the "I" for "intangibility," and the "P" for "perishability." Let us break these down, and then examine how each impacts the consumption process.

Simultaneity: Services are simultaneously produced and consumed. For example, a masseur creates an experience as the consumer experiences it. Massages are not stored in a spa, waiting to be pulled off a shelf when a customer buys one. The interaction between the service provider and the client is key to the service being customized and therefore optimally pleasing to the customer. The fact that services unfold in real time has several implications for the consumption process, as discussed below.

Heterogeneity: Services are also said to be heterogeneous. The idea is that one customer's experience is likely to differ from that of another. One person may rave about a masseuse, whereas another might find that provider's techniques too light or too deep, and he/she may not leave the spa as satisfied as the first customer. This type of difference is due to customer variability; on a large scale, we would call these segment preference differences. Other differences are attributable to the service providers having different styles and talents. Still other sources of variability are more temporary – a customer or service

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

provider may be in a bad mood, and have an "off" day. As these examples illustrate, heterogeneity in customer experiences is largely attributable to the service encounter being an interpersonal interaction. Marketing managers responsible for services have a much greater challenge than those overseeing goods; people are harder to manage than machines. We will see shortly how marketers try to compensate in the consumption process.

Intangibility: Services are also said to be more intangible than goods. A 17-year-old girl who bought the book *Twilight* has a tangible proof of purchase. Another, who went to see the movie, has no visible evidence of the movie experience, except, perhaps, for swoony eyes. Many purchases are of course a mix of the tangible and intangible. Imagine an evening dining out – there exists both the tangible (food, restaurant appearance) and the intangible (attentiveness of the wait-staff, ambience). Most people's largest purchases – a house and a car – include the physical purchase of the dwelling or the vehicle, as well as less-tangible elements, such as financing, insurance, and maintenance. For those two purchases, the tangible might seem to dominate the intangible. For another large household purchase, a kid's education, the intangible seems to dominate the tangible. These examples notwithstanding, expensive purchases are not always those dominated by goods – divorces are services and they are also expensive. Health care is a service and it can be costly. We now discuss what marketers make of intangibility.

Perishability: The final piece of the SHIP acronym is perishability, and this aspect of services derives in part from the first, the simultaneity of production and consumption. The potential service provision during lull times cannot be realized during peak times; it has perished. For example, the airplane leaves the ground with some empty seats that can never be regained, or a tax consultant's hours during a slow month cannot be recaptured in early April. Now let us blend these service dimensions with the consumption process.

What are the Implications for the Consumption Process?

In this section, we take the properties of simultaneity, heterogeneity, intangibility, and perishability and consider their effects on the service encounter. Recall the basic consumer behavior process: needs identification, information search, choice and purchase, and postpurchase reflection. We will treat the process in temporal order.

Needs Identification. While some services require special skills (e.g., tax attorney, psychiatrist), there are many services that customers can provide for themselves, should they choose to do so (e.g., lawn-mowing, ride to the airport). When customers choose to outsource and buy a service, there must be a reason why paying another to execute the service is superior to doing it oneself. The trade-off is usually some combination of time, expertise, and convenience. For example, most of us could mow our lawns, if we wanted to: we would start up a mower, push the machine back and forth, and so on. However, most of us are busy enough that we would assess a trade-off of time and value, and conclude that it is worth it to hire someone to take care of the landscaping.

Yet while it is true that we seem to be getting busier and busier, there is a seemingly counter trend, and that is the growth of self-services (Meuter *et al*., 2005). Many grocery chains offer self-service checkouts. At the outset, checking oneself out might seem to take more time than letting someone else scan the items, but when the lines for the checkers are long, the same principle of time-famine motivates customers to do the work themselves – it is faster. There is no price discount (or premium) for self-checkout. There are also no superior skills required of the checker than the customer. Hence, the self-service checkout is altogether more convenient.

As the effects of information technology seem omnipresent, here too, smart software can begin to simulate the expertise that a regular customer lacks and who would have traditionally sought a professional. There are a variety of software packages that facilitate our lives: some prompt us to calculate proper taxes, others assist in writing wills, other software functions as a dietitian, and still other programs serve as tutors when learning foreign languages.

The challenge in the future will be to determine those purchases for which we will continue to buy a service provider's time and expertise, and those for which we will enjoy the consumption of the activity itself. Marketers for service providers will have the challenge of demonstrating even greater value in the time and money saved, and the additional expertise and value-added elements that the service provider may bring to the customer.

Information Search. For reasons that will become clear shortly, services marketers have always made much of the expectations that customers hold when they enter the service encounter. Marketers try to shape expectations through various communications, for example, advertising, public relations, and information posted online. Yet unlike shopping for a laptop or automobile, where we might seek comparisons among brands as conducted and reported by experts, a good part of services are experiential. That services are experiential is attributable in part to the intangibility quality, an essence that is difficult for an expert to convey.

The challenge to marketers is to take the intangibility, which is difficult to grasp by customers because it is abstract and ethereal, and make it more concrete. As a result, marketers are encouraged to use symbols in their advertising and logos to express their benefits more tangibly (e.g., "You're in 'good hands' with AllState"; "Like a 'good neighbor,' State Farm is there"). The abstraction also leads customers to believe that the purchase might be somewhat risky, as they do not know quite what to expect. Accordingly, some marketers have demonstrated success in showing customers videos that preview the service they are about to experience so as to help them understand more precisely what to expect (Bitner *et al*., 1997).

Customers who have experience with a particular service provider will naturally have information from which they derive expectations for future interactions. Lacking direct experience, customers are thought to rely on word-of-mouth from trusted sources, such as friends or coworkers. Indeed, services marketers have long posited the very likely greater importance of word-of-mouth for services than for the purchase of goods. As online social networks thrive, we can expect this form of tailored communication to only grow in strength. Marketers have not yet come upon the winning formula for monetizing a presence in such spaces, but with time, no doubt, will do so.

Another prepurchase phenomenon in services that receives much consideration is queuing, the one tool in a marketer's arsenal that approximates the inventorying of goods. While a movie theater might not be able to seat more than 200 patrons on a Friday night for an opening of some new, wildly anticipated movie, it can encourage moviegoers to attend at another showing, by making another time slot more desirable, such as by lowering prices. Restaurants similarly take reservations to try to manage fluctuating demand. Yield management is a near science at the airlines, though not without error, as overbooking experiences will attest. In fact, marketers look to levers such as price to increase or decrease demand, so as to more optimally match supply, in the provision of many kinds of services (Desiraju and Shugan, 1999).

Choice and Purchase. The purchase of a service is far more complicated than the purchase of most goods. The complications are first due to the heterogeneity, due to the fact that the service encounter is frequently an interaction between two people, the customer and service provider, both of whom can introduce noise or error into the system. There is not much that a marketer can do to control the behavior of a customer, but the behavior of the front-line employee serving that customer leads marketing to a nexus with human relations. To ensure smooth service encounters, marketers have emphasized the importance of supervisors and companies being selective in hiring service providers and in training them (Heskett, Sasser, and Schlesinger, 1997), empowering them (Schneider and Bowen, 1995), and keeping them motivated (Rust and Chung, 2006).

The complications of services marketing purchases are also due to the aspect of services being simultaneously produced and consumed. One implication of the fact that services unfold in real time is that the customer's evaluation of the service – the service experience, the front-line service provider, or the service firm brand – is comprised of many "moments of truth" (Carlzon, 1989). All these elements, whether large and central (e.g., a dental exam),

or smaller and seemingly peripheral (e.g., the availability of parking near the dentist's office), factor into the customer's overall assessment of the service encounter. To manage the multitude of process issues, services marketers have encouraged the use of flow charts to explicate the consumer experience down to the last detail, as well as the behind-the-scenes operational elements which facilitate a smooth customer service flow (Lovelock and Wirtz, 2006; Fisk, Grove, and John, 2007; Bitner, Ostrom, and Morgan, 2008). Note too that the real-time process can be potentially advantageous – it can allow for the customization of the transaction for superior service, or for the recovery of impending service failure for at least minimally acceptable levels of service quality.

As brand building is all the rage, it is important to see the implications of the heterogeneity and simultaneity on this marketing effort. It is difficult to build a brand, much less an excellent brand, when it is difficult to assure the consistency of the quality of the purchase experience. In services marketing, one form of brand building is the distribution channel of the franchise system. In franchising, the attempt is to sell a service system that is as standardized, replicable, and as consistent as possible. Some succeed – those with a large portion of tangible character (e.g., some hotel or restaurant chains), or with heavy assistance of information technology (e.g., tax assistance offices). However, while brand building might be an admirable goal for such straightforward services, it would likely be more difficult to brand a professional service. In the professions, one touts the reputation, that is, the brand, and hopes that all the service providers deliver at the heights implied and expected. If the members of the profession have been hired selectively and well trained, perhaps the service can be seamless and of consistently high quality.

Postpurchase Reflection. Services marketing adheres rather uniformly to the model known variously as the *disconfirmation paradigm* of customer satisfaction or the *gap model* of service quality (Oliver, 2009; Zeithaml, Parasuraman, and Berry, 2009). In this model, a customer is thought to come to the service encounter with certain expectations, and judges the quality of the encounter against those expectations. If the customer's expectations are met, it is predicted that the customer is satisfied. If the experience is much better than the expectations, it is predicted that the customer is delighted. If the experience does not hold to the expectations, the customer is expected to be dissatisfied.

It is in this evaluative phase that we see why marketers pay attention to expectations. Some admonish to "manage expectations" (downward), so that the service provider can meet or exceed them. However, downward management is counter to the marketing impulse of bragging in advertising.

The comparative judgment model got attention very early as services marketers struggled with the implications of services being intangible, and therefore quality being difficult, or at least subjective to evaluate. Marketers overseeing goods of many kinds, from consumer packaged goods to electronics to durables, rode a wave in the 1980s of total quality management, which specified, among other things, accuracies in processes and outcomes. Manufacturers of corn flakes can say they want 10 ounces of flakes per box, and if they set their machinery properly, that is what they will get. Marketers responsible for a high quality service encounter struggle with such standards. What should be measured and managed that serve as cues to good quality? In the late 1990s, many companies would brag, "We answer our customer service calls within three rings," which of course begs the question, "What is the quality of the customer service provided once the phone is picked up?" As a result, the evaluation of customer satisfaction or service quality remains largely subjective, perhaps as it should be, if we marketers truly believe customers should have the last word.

As the consumption process of the service encounter is extended over time, it is natural to take the evaluation and generalize it as well. Loyalty systems thrive in services, at least in those sectors where switching costs are not high. Whether it is a free coffee for every 10th purchase, upgrades for sufficient air miles, shorter queuing systems for preferred insurance packages, or discounts for the use of retail cards, services marketers have embraced loyalty programs. In turn, when the resulting information capture is used creatively, true customer relationship management can further enhance

the customer's share of heart and wallet dedicated to the particular service provider.

Conclusion

Consumers are not thought of as behaving all that differently in a service encounter compared to how they behave when purchasing a good. However, the service encounter itself is a more complicated purchase. Accordingly, as the services marketing literature matured, the research and writings gave opportunity to reflect on new issues.

As stated at the outset, most purchases have some element of goods and some element of services transacted. The question is a matter of the proportions of each. As a result, the theories and research findings of the services marketing scholars are relevant and very broadly applicable.

Bibliography

Bitner, M.J., Ostrom, A.L., and Morgan, F.N. (2008) Service blueprinting: a practical technique for service innovation. *California Management Review*, **50** (3), 66–94.

Bitner, M.J., Faranda, W.T., Hubbert, A.R., and Zeithaml, V.A. (1997) Customer contributions and roles in service delivery. *International Journal of Service Industry Management*, **8** (3), 193–205.

Carlzon, J. (1989) *Moments of Truth*, Harper & Row, New York.

Desiraju, R. and Shugan, S.M. (1999) Strategic service pricing and yield management. *Journal of Marketing*, **63** (1), 44–56.

Fisk, R.P., Grove, S.J., and John, J. (2007) *Interactive Services Marketing*, Southwestern, Dallas, TX.

Heskett, J.L., Sasser, W.E. Jr., Schlesinger, and L.A. (1997) *The Service Profit Chain*, Free Press, New York.

Lovelock, C. and Wirtz, J. (2006) *Services Marketing*, 6th edn, Prentice-Hall, New York.

Meuter, M.L., Bitner, M.J., Ostrom, A.L., and Brown, S.W. (2005) Choosing among alternative delivery modes: an investigation of customer trial of self-service technologies. *Journal of Marketing*, **85**, 61–83.

Oliver, R.L. (2009) *Satisfaction: A Behavioral Perspective on the Consumer*, 2nd edn, McGraw-Hill, New York.

Rust, R.T. and Chung, T.S. (2006) Marketing models of service and relationships. *Marketing Science*, **25** (6), 560–580.

Schneider, B. and Bowen, D.E. (1995) *Winning the Service Game*, Harvard Business School, Boston, MA.

Zeithaml, V.A., Parasuraman, A., and Berry, L.L. (2009) *Delivering Quality Service*, Free Press, New York.

consumer aspects of international marketing

Claudiu V. Dimofte

The Global Consumer

The globalization of markets has long defined the research agenda of economists and marketers alike. Yet the perspective that each group tends to employ when approaching the topic varies to a large extent: the former are mostly concerned with the international flows of capital, labor, and goods that imply a large degree of cross-market homogeneity, whereas the latter generally assume cultural heterogeneity and look for locally defined cultural determinants of consumption behavior. On the one hand, marketplace evidence suggests that specific consumption patterns across the globe do show some convergence, such that market segments from culturally and geographically distant nations are developing similar tastes and product preferences. This phenomenon is at the core of what GLOBAL MARKETING STRATEGY has termed the *universal segment* (or *intermarket*) approach to market segmentation: positioning the firm's offerings toward groups of consumers around the world who display similar needs and look for the same product benefits, regardless of location (see Toyota Scion vehicles and MTV programming targeting Generation Y consumers internationally). Furthermore, the ubiquity of the Internet at the center of many nations' social and technological environments allows marketing information to easily diffuse and spill over, to create more uniform consumer knowledge and interests, and to generate relatively consistent brand preferences across borders, as the law of communicating vessels would predict.

On the other hand, the alternative view to this single world economy is the perception of multiple world markets, each described by idiosyncrasies that are deeply rooted in cultural heritage, as determined by local knowledge, beliefs, customs, and so forth. Global marketing textbooks provide a multitude of approaches for segmenting such markets, under the implicit assumption that attempting to standardize one product across national boundaries described by so much cultural variance is practically impossible. For anyone who has ever traveled to a foreign land and has experienced the at once wonder and shock associated with dramatically different people, foods, or customs (ever had a balut egg in the Philippines or exchanged business cards in Japan?), this perspective has immediate resonance. To better understand what global consumers may be about and how marketers can appropriately describe and predict their behavior, it is useful to begin by addressing the topic of consumer culture (*see* SOCIETY, CULTURE, AND GLOBAL CONSUMER CULTURE) in more detail.

Consumer culture. Defining the *culture* construct is no easy theoretical endeavor (*see* SOCIETY, CULTURE, AND GLOBAL CONSUMER CULTURE). In short, culture consists of a society's complex set of beliefs, customs, and other specific forces that shape individuals' perceptions and behaviors and is thought to be a key characteristic of a national environment that produces systematic differences in behavior across borders (Steenkamp, 2001). Whereas some have suggested that clusters of countries may display numerous cultural commonalities and could thus be construed as regional metacultures (Steenkamp, 2001), dealing with the idiosyncrasies of national cultures is a more intriguing conceptual endeavor. Looking for the dimensions underlying cultural specificity has thus been the main concern of cross-cultural research, in an attempt to bring some uniformity and measurement structure to a highly diverse global environment.

The earliest and arguably the most influential account of cultural variation and its impact on business is that of Hofstede (1980). His framework discusses five dimensions: individualism/collectivism (addressing the individual vs group relationships), power distance (looking at power inequality in terms of hierarchical social layers), masculinity/femininity (dealing with the social implications of gender, such as a focus on success vs quality of life), uncertainty avoidance (handling economic and social uncertainty), and long-term orientation (displaying a cognitive focus on short- vs long-term temporal horizons). The empirical study underlying his research was highly comprehensive and robust, employing over 60 000 respondents across 70 countries. Each of these nations was assigned an index on

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

each dimension, which was then linked with a variety of societal variables (geo-demographics, political, economical, etc.) – an operationalization that no other research has matched.

Relative to other frameworks, Hofstede's (1980) work displays high relevance to a variety of global business contexts (including international marketing), while maintaining a high level of conceptual convergent validity. That said, Schwartz (1994) provides an alternative perspective that – unlike Hofstede's (1980) focus on social dimensions – takes a more individual-level look at specific cultural values. Within a comparably large sample (over 60 000 respondents across 64 nations), Schwartz (1994) uncovers seven universally relevant cultural domains: harmony, egalitarianism, intellectual autonomy, affective autonomy, mastery, hierarchy, and embeddedness. Owing to their complementarity, combining the Hofstede (1980) and Schwartz (1994) frameworks has been suggested as a way to uncover clusters of more manageable regional cultures of interest to marketers (Steenkamp, 2001).

The immediate practical relevance of the cultural specificity that these frameworks address for the marketing efforts of any firm can be observed even within the confines of a particular nation, as in the proverbial cultural melting pot represented by the United States. Simply thinking about the different consumption patterns that the various US ethnic groups display can easily illustrate this point. Reading the grocery shopping lists of families of consumers in the African-American, Hispanic, or Asian ethnic groups will probably reveal very different items, flavors, and quantities. And watching television programming (as well as the associated advertising) on NBC relative to Telefutura or the BET channel will most likely entail highly distinct experiences. The existence of these microcultures (Steenkamp, 2001) contoured along ethnic or other dimensions highlights the importance of studying cultural heterogeneity through a more fine-tuned lens.

Another relevant stream of research addresses what social psychologists have termed individual difference variables describing consumers. In an international marketing context, one must be aware of a variety of social or subjective norms that influence how consumers in a nation perceive or respond to (products from) specific countries (home or foreign). For example, some consumers display a tendency to denigrate foreign products and favor domestic ones, a behavior that renders them ethnocentric with respect to foreign products (*see* BASE OF THE PYRAMID MARKETS: CULTURE INSIGHTS AND MARKETING IMPLICATIONS). A related construct is that of CONSUMER ANIMOSITY, such as that describing reactions of some American consumers to French-made products after this nation's opposition to the 2003 invasion of Iraq (remember those *freedom* fries?). On the opposite end of the spectrum, there are individuals who show a consistent interest in other cultures and global issues or even preference for foreign products. These consumers are generally considered high on worldmindedness (*see* CONSUMER WORLD-MINDEDNESS), affinity (*see* CONSUMER AFFINITY CONSTRUCT), or cosmopolitanism.

If cultural and even individual differences are so prevalent both across nations and within them, what is left of the claim that today's consumer is a global consumer? Keillor, D'Amico and Horton (2001) undertook survey research with consumer samples from the United States, France, and Malaysia to assess the extent to which individuals with such culturally and economically distinct backgrounds would display similar tendencies related to specific aspects of the psychology of consumption. When juxtaposed across four relevant constructs (i.e., national identity, ethnocentrism, social-desirability bias, and response to consumer influence sources), the differences were not remarkable, suggesting that consumers around the globe may indeed become more homogenous in terms of their consumption behaviors. Nevertheless, although a global consumer may be emerging, cultural differences across nations are still highly relevant to international marketing researchers. The next section provides specific considerations in support of this point.

International marketing research. One of the classic debates in the academic field in terms of dealing with the cultural differences that underlie the consumption behavior of

individuals across nations is that between the emic and etic perspectives. Paralleling the earlier idiosyncratic–generic dichotomy related to the globalization of consumers and consumption patterns, this debate refers to the choice of theoretical frameworks to be employed in examining cultural differences. The *emic* view argues that such frameworks need to be culture-specific and operationalized via indigenous research techniques and instruments. Alternatively, the *etic* perspective favors the use of already established, universal theoretical frameworks as benchmarks for research pursued in any specific culture (Maheswaran and Shavitt, 2000).[1] A specific charge levied against the latter approach is that too often this established framework is developed by US researchers who simply attempt to apply it in novel cultural contexts, with little adaptation (Sekaran, 1983; Steenkamp, 2001). A potential solution to the emic–etic debate was proposed by Douglas and Craig (2006) in the shape of two iterative research approaches: the *adapted etic* (i.e., explicitly adapting the conceptual model of a base culture to the differences of new research contexts) and the *linked emic* model (i.e., taking the local context as the simultaneous starting point, at multiple research sites).

A related concern for international marketing researchers (academics and practitioners alike) is that of achieving measurement equivalence when it comes to employing specific scales in cross-cultural contexts. Therefore, it is important that not only conceptual equivalence of the target construct(s) be ensured (in particular within an emic theoretical approach) but that instrument equivalence (e.g., in terms of the item and scalar properties of the measure) be achieved as well (Maheswaran and Shavitt, 2000). In terms of construct operationalizations, method and item bias are often a concern in multicountry marketing research, as are the choice and relevance of the particular unit of analysis employed (Douglas and Craig, 2006). Finally, cross-cultural research is also affected by more practical methodological problems, such as the sampling, timing, and specific choice of statistical analyses underlying data collection and study across nations (Sekaran, 1983).

It should be noted that the field of international marketing is a broad one and includes theoretical accounts from multiple disciplines. Whereas the focus of this article is on consumer-level variables that borrow conceptually from social and cognitive psychology, there are complementary perspectives such as the theory of the international firm (e.g., global marketing strategy, *see* GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES, and the management of international operations) and the economics of global markets (e.g., international pricing, *see* INTERNATIONAL PRICING OBJECTIVES AND STRATEGIES, and distribution INTERNATIONAL MARKETING CHANNELS) which require similar concern with conceptualization, operationalization, and measurement. The next section bridges the global consumer–global firm dyad by addressing recent literature findings on consumer response to global brands (*see* GLOBAL BRANDING: THREE KEYS FOR GLOBAL BRAND SUCCESS).

Consumer Response to Global Brands

As is the case with the *culture* construct, there is also considerable ambivalence in the international marketing literature in terms of what a *global brand* actually represents. Much of the practitioner view on the issue is simply limited to the international, multimarket reach and profit potential that a brand accrues. For example, the Interbrand agency – which compiles its annual top 100 brands list for publication in Business Week – looks in its valuation at a brand's future earnings potential, with the expectation that a certain percentage of its sales come from each market in which it operates. The rival Millward Brown agency and its similar rankings published in the Financial Times require that a brand be present in at least seven countries before it is deemed global; yet by its own admission only 3% of its more than 10 000-brand database achieve this standard.

In terms of the academic research on the topic, the problem of operationalization may be one reason for the lack of a universally accepted definition of the global brand. However, the dimensions that underlie the global brand construct have been researched in a handful of articles, including Dimofte, Johansson and Ronkainen's (2008) work that identified five

latent factors: *reach*, *aspiration*, *low risk*, *ethics*, and *standardization*. The authors show that consumers perceive global brands to be widely available, well recognized, standardized, more powerful, more cosmopolitan, and subject to more stringent social responsibility standards than other, more local brands. However, other brand attributes such as higher quality are – perhaps somewhat surprisingly – not necessarily associated a priori with global brands.

Combining the practitioner and academic perspectives may bring into focus a clearer, albeit still imperfect, understanding of what the global brand is about. Thus, as far as consumers, brand managers, and most academics are concerned, a global brand has a strong presence across multiple national markets, which associates it with universal recognition, the promise of long-term financial survival, and the burden of higher social responsibility standards governing its corporate behavior. Yet, the largely positive image inherent to this global brand definition is not universal. Literature suggests that there are important differences in terms of the specific valence that these defining global brand attributes acquire for consumers in the developing versus the developed world. Simply put, whereas the former aspire to the idealistic qualities embodied by a global brand, the latter are significantly less impressed. The next section addresses this dichotomy in more detail.

Global brands in developing versus developed nations. Alden, Steenkamp and Batra (1999) argue that globally positioned brands might work better in markets that are characterized by lower levels of economic development, since "consumers in these markets may admire the 'economic center' and believe that [. . .] ownership of brands from the West increases the owner's status" (Alden, Steenkamp and Batra, 1999, p. 84). Subsequent work by Batra *et al.* (2000) finds similar evidence that consumers in developing nations perceive nonlocal brands as preferable to local brands.

Holt, Quelch and Taylor (2004) uncover four basic dimensions of global brands *quality signal*, *global myth*, *social responsibility*, and *American values*. Averaging across responses from 12 countries (including the United States), their results also suggest that developing nations consumers aspire to the greater global community and in the process downgrade their own brands relative to global brands.

While the evidence of the attractiveness that global brands entail for consumers in developing nations is consistent in the international marketing literature, things are less clear cut when it comes to consumers in developed nations. On the one hand, Steenkamp, Batra and Alden (2003) investigate US and Korean samples and find that perceived globality positively impacts both perceived brand quality and prestige, including in the United States. On the other hand, Johansson and Ronkainen (2005) analyze the top 150 brands in terms of brand strength from the large cross-country and multiyear database underlying Landor's Brand Asset Valuator (BAV) model. In a survey of these brands across eight developed countries, the authors find that although global brands are held in high esteem, they are not necessarily associated with high levels of quality or other desirable brand attributes. There is also some evidence suggesting that local brands still command deeper customer loyalty than global brands in developed countries. For example, empirical analyses of Young & Rubicam data on brands from the European marketplace show that in such mature markets local brands often do better with consumers than global brands.

To understand this international dichotomy of responses to global brands, Dimofte, Johansson and Bagozzi (2010) evaluate ethnic consumer groups in the United States. Reflecting the aspirational dimension inherent to global brands for consumers in developed nations, the historically more economically disadvantaged African-American and Hispanic market segments are shown to associate global brands with higher product quality, more social status and prestige, as well as superior style. Yet, Caucasian consumers do not find global brands to be any more exciting than other brands, and – although they do perceive them to imply higher social status – they do not think their product quality is necessarily higher, mimicking the findings in developed nations (Dimofte, Johansson and Bagozzi, 2010). However, as a conceptual check, the relationships in a structural equation model show similarity across

ethnic groups in the way the global brand associations link up to consumer attitudes and purchase behavior. Those with more positive perceptions of global brands also have more positive attitudes overall and, as one would expect, show higher purchase rates of global brands. Although Caucasians' perceptions and attitudes tend to be less favorable than those of African-American and Hispanic consumers, they nonetheless buy global brands at the same rate as non-Caucasians (Dimofte, Johansson and Bagozzi, 2010).

Other global branding research findings. The extensive literature on global branding cannot truly be done justice within the limiting confines of a short article, so the final part of this section presents selected findings on the topic from two lines of consumer research – both classic (the country-of-origin effect) and more recent (automatic consumer response to global brands).

A firm's marketing communications, strategic marketplace behavior, and perceived competitive performance create in consumers' minds a particular brand personality and a variety of related brand associations – in short, (*see* PERCEPTION OF BRAND EQUITY). One of these brand associations that is particularly relevant to international marketing researchers involves consumer perceptions of a brand's country of origin (*see* "COUNTRY OF ORIGIN" AS BRAND ELEMENT). The location where a product is made prompts immediate country associations (i.e., a particular country image) that consumers generally incorporate into their brand evaluations and oftentimes produce cognitive biases in terms of attitudes and choice. Consumers tend to evaluate products more favorably when they originate in countries that benefit from a positive image (i.e., stereotypical perceptions describing the likely quality and performance of locally made products).[2] Most of us would likely agree that given a choice and in the absence of price constraints we would prefer a Swiss watch, a French perfume, or a German car. At the same time, similar negative stereotypical perceptions (created or reinforced by significant press coverage) have recently produced an avoidance behavior among many consumers toward products originating in China.

The country-of-origin effect is robust enough to emerge in studies that employ single-cue and multiple-cue product stimuli, within and between-subjects designs, and US and non-US samples. There have been two major cognitive process explanations proposed to underlie this effect. Depending on a consumer's familiarity with a country's products, country image can operate as either an inferential halo or a summary construct. If a consumer is unfamiliar, country image is used as a halo that impacts attribute ratings and, indirectly, brand attitudes (e.g., for Korean automobiles, quality inferences are made based on one's perception of Korean-made things in general). If the consumer is familiar with the country's products, the country image becomes a summarizing construct for the product attributes, directly influencing brand attitudes (e.g., for Japanese electronics, the fact that they are made in Japan is a proxy for high quality).

An interesting cognitive bias related to the country-of-origin effect is presented by Leclerc, Schmitt and Dubé (1994) in their research on consumer response to brand names that merely suggest a specific national origin. The authors show that product evaluations change such that, for example, more hedonic perceptions follow a French pronunciation relative to an English one for the very same brand name (Leclerc, Schmitt and Dubé, 1994). The fact that merely altering the pronunciation produces biased perceptions of brand attributes (as mediated by country-of-origin inferences) attests to the strength and ubiquity of the effect in daily consumption. This is an important finding, because other research has shown that a product's country of origin heightens or lowers relevant consumer expectations (depending on the direction of the associated country stereotype), with direct implications for subsequent evaluative standards and customer satisfaction (*see* CUSTOMER SATISFACTION/DISSATISFACTION).

Despite the fact that the country-of-origin effect has been widely demonstrated, consumers' reliance on country image when forming attitudes and making choices is generally not acknowledged in self-reports. It has also been shown that country stereotypes can be spontaneously activated by the mere contextual

presence of country-of-origin information, with subsequent measurable impact on attribute and product evaluations without consumers' intention or control. A similar finding in terms of American consumers' response to global brands shows that they explicitly deny the importance of brand globality in choice, yet clearly favor them in terms of subsequent behavioral measures (Dimofte, Johansson and Ronkainen, 2008).

The lack of acknowledgement in self-reports could have two possible explanations. On the one hand, it may simply be a matter of conscious self-presentation bias, driven by a strategic desire to appear open-minded and sophisticated when responding to international marketing surveys. On the other hand, it may be that consumers are truly unaware of the biases they hold, and as such unconsciously rely on information such as a brand's country of origin or global nature to inform their subsequent judgments (*see* IMPLICIT CONSUMER COGNITION). To evaluate these alternative explanations in the case of global brands, Dimofte, Johansson and Ronkainen (2008) employed the Implicit Association Test (IAT).[3] The results showed that, compared to local brands, global brands were more closely associated with adjectives such as *ideal* and *desirable*, whereas local brands, by contrast, were seen as more concrete and mundane. The IAT responses did not correlate with self-reported explicit attitudes toward global brands, such that even consumers who were explicitly against globality displayed the same implicit preference for these brands.

Conclusions

The emergence of global consumption patterns and, at the same time, the complexity involved in understanding the still remaining national differences attest to the importance of research in international marketing. As reviewed in this article, the field is indissolubly linked to the topic of cross-cultural consumer behavior and its applications.

When considering the relevant conceptual and theoretical perspectives employed in research in international marketing, the emic–etic juxtaposition should be carefully assessed in order to ensure the validity of subsequent findings and their relevance to corporate strategy. At the same time, when thinking about the practical aspects of assessing and satisfying consumer needs across the globe, issues such as the appropriate operationalization of constructs and the achievement of measurement equivalence across multinational samples should be at the center of international marketing research. Yet, despite these cross-national cultural differences, a global consumer culture may be emerging, and the field seems well positioned to address it.

Finally, the extensive literature on relevant topics such as the country-of-origin effect and global branding is informative to both academics and practitioners in suggesting that, despite local specifics, the global consumers' behavior is governed by universal social-psychological processes and cognitive biases. Understanding these underlying factors will leave both consumers and firms better off.

Endnotes

[1] Others have argued that the international marketing field may benefit from employing established theoretical frameworks that originate in other academic areas.

[2] The level of fit between particular products and countries of origin underlies the product ethnicity construct (*see* PRODUCT ETHNICITY).

[3] The IAT is designed to measure the strength of automatic associations between mental representations of concepts in memory. It requires the rapid categorization of various stimulus objects, such that easier pairings (i.e., faster response latencies) are interpreted as being more strongly associated in memory than more difficult pairings (i.e., slower response latencies).

Bibliography

Alden, D.L., Steenkamp, J.B. and Batra, R. (1999) Brand positioning through advertising in Asia, North America and Europe: the role of global consumer culture. *Journal of Marketing*, **63**, 75–87.

Batra, R., Ramaswamy, V., Alden, D.L. *et al.* (2000) Effects of brand local and nonlocal origin on consumer attitudes in developing countries. *Journal of Consumer Psychology*, **9**, 83–95.

Dimofte, C.V., Johansson, J.K. and Bagozzi, R.P. (2010) Global brands in America: how consumer ethnicity mediates the global brand effect. *Journal of International Marketing*, **18** (3) in press.

Dimofte, C.V., Johansson, J.K. and Ronkainen, I.A. (2008) Cognitive and affective reactions of American consumers to global brands. *Journal of International Marketing*, **16**, 115–137.

Douglas, S.P. and Craig, S.C. (2006) On improving the conceptual foundations of international marketing research. *Journal of International Marketing*, **14**, 1–22.

Hofstede, G. (1980) *Culture's Consequences: International Differences in Work-Related Values*, Sage, Beverly Hills.

Holt, D.B., Quelch, J.A. and Taylor, E.L. (2004) How global brands compete. *Harvard Business Review*, **82**, 68–81.

Johansson, J.K. and Ronkainen, I.A. (2005) The esteem of global brands. *Journal of Brand Management*, **12**, 339–354.

Keillor, B.D., D'Amico, M. and Horton, V. (2001) Global consumer tendencies. *Psychology and Marketing*, **18** (1), 1–19.

LeClerc, F., Schmitt, B.H. and Dubé, L. (1994) Foreign branding and its effect on product perceptions and attitudes. *Journal of Marketing Research*, **31**, 263–270.

Maheswaran, D. and Shavitt, S. (2000) Issues and new directions in global consumer psychology. *Journal of Consumer Psychology*, **9**, 59–66.

Schwartz, S.H. (1994) Beyond individualism/collectivism: new cultural dimensions of value, in *Individualism and Collectivism: Theory, Method, and Applications* (eds U. Kim, H.C. Triandis, C. Kagitcibasi *et al.*) Sage, Thousand Oaks.

Sekaran, U. (1983) Methodological and theoretical issues and advancements in cross-cultural research. *Journal of International Business Studies*, **14**, 61–73.

Steenkamp, J.B. (2001) The role of national culture in international marketing research. *International Marketing Review*, **18**, 30–44.

Steenkamp, J.B., Batra, R. and Alden, D.L. (2003) How perceived brand globalness creates brand value. *Journal of International Business Studies*, **34**, 53–65.

consumers' need for uniqueness

Ayalla A. Ruvio

Consumers often use possessions to differentiate themselves from other consumers. This tendency of using possessions to project a unique identity has been conceptualized as consumers' need for uniqueness (CNFU). The theoretical basis of CNFU is rooted in the more general theory of the need for uniqueness (NFU), which focuses on people's perceptions of and reactions to their similarity to others (Snyder and Fromkin, 1980). NFU theory postulates that individuals continuously evaluate their degree of similarity or dissimilarity to others and act on such evaluations. Individuals will normally attempt to maintain moderate uniqueness from others and will actively avoid being too similar or too different from others. High levels of similarity or dissimilarity will result in unpleasant feelings, reducing the individuals' self-esteem (Fromkin, 1972), and will trigger emotional and behavioral reactions. However, the magnitude of such reactions will vary across individuals and situations and will be determined by the individuals' level of NFU. The stronger it is, the greater the individuals' sensitivity to similarity and the more different from others they will want to be (Fromkin, 1972). However, the desire for uniqueness does not ignore social norms and is constrained by the need for social assimilation and approval (Snyder and Fromkin, 1980).

Building on the notion that possessions are an extension of one's POSSESSIONS AND SELF, CNFU reflects the individual's propensity for acquiring, using, and disposing of possessions in order to construct and maintain a unique self and social image (Tian, Bearden and Hunter, 2001). Tian, Bearden and Hunter (2001) conceptualized CNFU as a behavioral construct with three facets. Creative choice counterconformity captures the individual's tendency to use socially acceptable products and brands in order to create a distinct personal image (Snyder and Fromkin, 1980). Unpopular choice counterconformity postulates that individuals who seek to differentiate themselves from others may do so by using products that are not entirely within group norms, and will risk social disapproval (Tian, Bearden and Hunter 2001). Finally, avoidance of similarity reflects the individual's intentional avoidance of popular and commonly used products and brands. Such individuals lose interest in, avoid purchasing, or discontinue using those brands when they become commonplace (Tian, Bearden and Hunter, 2001). Tian and her colleagues (2001) developed a 31-item scale to measure the enduring trait of CNFU. A shorter 12-item version was later validated by Ruvio, Shoham and Makovec-Brencic (2008) in different cultural settings.

Empirical studies established that individuals will tend to display their uniqueness more in identity-relevant domains and in situations that pose a threat to their sense of distinctiveness (Berger and Heath, 2007). CNFU was also found to be related to a variety of consumption behaviors such as shopping innovativeness (*see* CONSUMER INNOVATIVENESS), opinion leadership and market mavenism, optimum stimulation level, status consumption, individualism, choice of shopping venues, preferences for unique exterior product design, and preferences for scarce and customized products (e.g. Clark and Goldsmith, 2005; Ruvio, Shoham and Makovec-Brencic, 2008; Tian, Bearden and Hunter, 2001).

Bibliography

Berger, J. and Heath, C. (2007) Where consumers diverge from others: identity-signaling and product domains. *Journal of Consumer Research*, **34**, 121–134.

Clark, R.A. and Goldsmith, R.E. (2005) Market Mavens: psychological influences. *Psychology and Marketing*, **22** (4), 289–312.

Fromkin, H.L. (1972) Feelings of interpersonal undistinctiveness: an unpleasant affective state. *Journal of Experimental Research in Personality*, **6** (2-3), 178–182.

Ruvio, A., Shoham, A. and Makovec-Brencic, M. (2008) Consumers' need for uniqueness: short-form scale development and cross-cultural validation. *International Marketing Review*, **25** (1), 33–53.

Snyder, C.R. and Fromkin, H.L. (1980) *Uniqueness: The Human Pursuit of Difference*, Plenum, New York.

Tian, K.T., Bearden, W.O. and Hunter, G.L. (2001) Consumers' need for uniqueness: scale development and validation. *Journal of Consumer Research*, **28** (3), 50–66.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

integrated marketing communication

Philip J. Kitchen and Inga Burgmann

Introduction

Integrated marketing communication (IMC) emerged during the late twentieth century and its importance has been growing ever since (Grove, Carlson, and Dorsch, 2002; Cornelissen, 2001; Hartley and Pickton, 1999). Owing to the impact of information technology, changes came about in the domains of marketing and marketing communications which led to the emergence of IMC (Kitchen *et al.*, 2004a; Phelps and Johnson, 1996; Duncan and Everett, 1993). The multiplication of media, demassification of consumer markets, and the value of the Internet in today's society are just three of the areas in which technological innovation has impacted (Pilotta *et al.*, 2004; Peltier, Schibrowsky, and Schultz, 2003; Reid, 2003; Lawrence, Garber, and Dotson, 2002; Fill, 2001; Low, 2000; Hutton, 1996). This in turn left marketers in a challenging and competitive environment, trying to fulfil customers wants and needs while also developing long-term relationships with them. IMC can help in creating coordinated and consistent messages across various channels of communication. Furthermore, the concept is especially valuable in that it places great emphasis on the importance of all stakeholder groups and, in particular, on customer loyalty, which can only be created through strategic relationship building (Jin, 2003/2004; Cornelissen, 2000; Eagle and Kitchen, 2000; Pickton and Hartley, 1998; Miller and Rose, 1994).

To date, academic research on IMC has been limited. The majority of empirical research has been conducted with advertising and PR agencies or companies located in the United States with a clear tendency toward quantitative methodologies (Jin, 2003/2004; Peltier, Schibrowsky, and Schultz, 2003; Kitchen and Schultz, 1999; Beard, 1996; Miller and Rose, 1994). Very few studies have concentrated on the business "client" perspectives within Europe (Cornelissen and Thorpe, 2001; Low, 2000; Kitchen and Schultz, 1999).

The following sections discuss IMC in greater depth. Firstly, the relevance of IMC as a concept is discussed. Subsequently, an analysis of IMC implementation is carried out with reference to past research. Then, barriers to integration and criticisms of IMC are addressed. Finally, concluding comments are proffered.

The Importance of IMC

This section focuses on the importance of IMC. First, the evolution of IMC and different IMC definitions are reviewed. A contemporary perspective on IMC is offered with a subsection flagging up the benefits of such an approach. Lastly, IMC's impact on promotional mix elements is presented.

The evolution of IMC. Integration, the attempt to present a consistent message across the available promotional mix elements has always been important to successful organizations even during the mid twentieth century. With the multiplication of media channels in the late twentieth and early twenty-first century, the integration and coordination of different messages aiming to portray a single and unique image to all stakeholder groups has become both more important and more difficult to achieve.

However, some researchers believe that the concept of IMC can be traced back to the 1970s (Cornelissen and Lock, 2000; van Riel, 1995). The first study on IMC was conducted by Caywood, Schultz, and Wang (1991b) at the end of the 1980s, while the first "conceptual ideas" were published in the book "Integrated Marketing Communications" by Schultz, Tannenbaum, and Lauterborn (1993). IMC advocates believe that its emergence was down to the context of media upheaval of that time, for example, digital TV and mobile phones, market environments, that is, increasing global competition and rapid technological developments, such as the personal computer (Kliatchko, 2005; Reid, 2003; Eagle and Kitchen, 2000; Griffin and Pasadeos, 1998; Bruhn, 1997/1998; Hutton, 1996).

Technology can affect IMC from two sides, that is, from the marketing and consumer perspectives (Kitchen *et al.*, 2004a; Schultz, 1993f). Today, integration is needed owing to globalization and the resulting interdependence between countries and marketplaces (Kitchen

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

et al., 2004a; Schultz, 1996b). Thus, corporate and brand managers need to coordinate the actions of their global and even national brand(s) with the aim of integrating elements of promotional mix.

A global marketplace which becomes more transitory through the Internet may lead to a customer-driven and focused marketing environment. In such an environment, technology can enhance marketing communication strategies, that is, both traditional advertising techniques and also new, unconventional marketing practices may be applied, such as database marketing, one-to-one communication, or marketing PR (McGrath, 2005a; Edelman, 2004; Gonring, 1994; Nowak and Phelps, 1994). The social networking website MySpace, for example, has more than 17 million visitors each month in the United States. Marketing communication innovators propose the exploitation of such networks as a means of establishing an ongoing dialogue with customers and prospects (Economist, 2006a). In conclusion, not only can technology help consumers to connect and communicate with each other but it can also act as a facilitator in establishing a relationship with individual consumers and companies.

Owing to disagreements about the emergence of IMC and the limited amount of research conducted with actual organizations, researchers have thus far been unable to agree upon a single definition of IMC. Kliatchko (2005), though, has examined the various definitions of IMC that have emerged over time and tried to explore their dimensions in the hope of unifying them. Table 1 is based on Kliatchko's (2005: 21) table of IMC definitions. So, his own definition and another recent definition by Schultz (2004b) have also been added.

As demonstrated in Table 1, during the early 1990s IMC was referred to as the *one sight, one sound* or *one voice* or the *seamless marketing communication* approach (Beard, 1997; Nowak and Phelps, 1994; Duncan and Everett, 1993). However, the increasing interest in the subject of IMC led researchers to reevaluate the concept and the "buzz words" were soon set aside (Grove, Carlson, and Dorsch, 2002; Lee, 2002; Fill, 2001; Hartley and Pickton, 1999; Phelps and Johnson, 1996). These words only hinted at the many applications of IMC and, therefore, new concepts were added to the earlier definitions of IMC.

Contrary to the implication of many of these "buzz words," IMC does not mean that an organization should only work with one message or with a single unifying brand. Rather, an integrated approach encourages managers to work with multiple targets and enables them to achieve integration of different brands, communication messages, and functions within one company. Thus, IMC has the potential to fundamentally change the meaning of marketing communications and may even be the next step in the evolution of marketing (Dewhirst and Davis, 2005; Kliatchko, 2005; Grove, Carlson, and Dorsch, 2002; Lee, 2002; Phelps and Johnson, 1996).

As in an early definition by Caywood, Schultz, and Wang (1991b) (see below), for many practitioners and researchers, IMC is a concept which strengthens the effects of promotional mix elements at the tactical level, if these elements are united (Schultz, 2006a; Cook, 2004; Naik and Raman, 2003; Fill, 2001; Cornelissen and Lock, 2000).

> [IMC is a] concept of marketing communications planning that recognizes the added value of a comprehensive plan that evaluates the strategic roles of a variety of communication disciplines – general advertising, direct response, sales promotion, and public relations – and combines these disciplines to provide clarity, consistency, and maximum communication impact (Caywood, Schultz, and Wang, 1991b: 2-3).

However, more recent definitions, such as the one presented by Schultz (2004a), add substantial value to old definitions and to the term IMC more generally. The concept is now viewed as a strategic instrument (Schultz, 2004b: 9).

> IMC is the concept and process of strategically managing audience-focused, channel-centerd, and result-driven brand communication programs over time (Kliatchko, 2005: 21).

The new definition seems more appropriate to the twenty-first century, as Schultz acknowledges IMC's importance as a business process and its value to both external and internal audiences. The definition provided by Schultz

Table 1 IMC definitions.

Author and Year	*Concepts Introduced*
Caywood, Schultz, and Wang (1991) and Caywood, Schultz, and Wang (1991b)	• Coordination and consistency of messages and communication channels (one sight, one sound) • Use of a variety of communication disciplines to work in synergy based on a comprehensive plan • IMC as a concept
Schultz (1991)	• Inclusion of consumers, prospects • Behavioral responses • Nurture relationship and customer loyalty • IMC as a process
Duncan and Everett (1993)	• Profitable relationships expanded audience scope from customers to other stakeholders
Nowak and Phelps (1994)	• Reinforced notions of consistency, coordination, and behavioral response
Schultz and Schultz (1998)	• Strategic business process • Expanded notion of brand communication • Measurability • Specified the multiple markets more explicitly, inclusive of external and internal audiences
Schultz (2004b) and American Marketing Association (2007)	• Strategic business process • Extensive brand communication • Evaluation and measurement • External and internal stakeholder groups • Long-term brand value focus
Kliatchko (2005)	• Process and concept • Audience-focused • Communication program • Result-driven

Source adapted from Kliatchko (2005: 21)

(2004b) has also been used by the American Marketing Association to define integrated brand communication (American Marketing Association, 2007). However, Kliatchko (2005) does not focus explicitly on individual stakeholder groups but rather refers to them as "audiences." In addition, his definition is not as explanatory in nature as Schultz's (2004b), as he underplays the results-driven characteristic of IMC and does not explicitly mention the importance of the long-term brand value and short-term financial returns. As noted by Kliatchko (2005); Schultz's (2004b) definition supplements other IMC definitions through its employment of the terms "business process," "evaluation," and "measurability".

Many researchers have noted that it may not be possible to agree upon a universal IMC definition, given the various interpretations of IMC and its different values in the academic and commercial spheres (Kliatchko, 2005; Phelps and Johnson, 1996; Stewart, 1996). A critical review of previous definitions of IMC and an assessment of current IMC literature reveals that researchers were able to reduce any IMC definition to five crucial attributes (Kitchen *et al.*, 2004a; Low, 2000):

1. The communication effort should be directed at consumers in order to affect behavior.
2. An outside-in approach should be utilized, that is, start with the customer first when developing a communication strategy.
3. A well-established relationship between the company and the customer is necessary.
4. To deliver a message correctly all communication activities should be included with contact points integrated into the strategy.
5. To create a competitive brand, coordination between the communication disciplines is needed.

A contemporary perspective of IMC. Owing to the rising demand in various product categories after World War II, the focus of marketers shifted to the product itself, making potential relationships with customers something of a side issue (Johnson and Schultz, 2004; Kitchen, 1999; Evans and Berman, 1987). During the 1950s, most organizations arranged their plans in line with their products and not according to customer wants and needs (Johnson and Schultz, 2004). Fuelled by dramatic changes in the marketplace and communications, retailers and other intermediaries have since gained in importance (Rosenbloom, 2004; Low and Mohr, 1999; Schultz, 1996b).

With competition increasing and with supply outstripping demand in most developed countries, a shift in power towards retailers and intermediaries has occurred. For this reason, building a long-term relationship with customers should be important to any kind of business in today's marketplace. Indeed, owing to the direct contact that retailers have with the end-consumers of manufactured goods, they should also know them better. Indeed, a powerful relationship exists between retailers, intermediaries, and manufacturers. (Reid, 2003; Low and Mohr, 1999; Schultz, 1996b). In an ideal world, the relationship between retailer, intermediary, and manufacturer would be interdependent, if not to say integrated, so that all parties could create customer-orientated communication messages, and essentially so that the manufacturer could develop customer-specific products (Pickton and Broderick, 2005; Schultz and Schultz, 2003; Schultz, 1993c).

Communication between customer and manufacturer used to be linear or one-way, that is, from the marketer to the consumer (Schultz, 1993c). But the customer of today has knowledge about the marketplace and has no need to wait for messages and information from the organization. The customer of today makes informed decisions and will make demands and influence advertisements or any other information received. Marketers must respond to these changes. The marketer needs to develop communication plans with the consumer as the starting point, later working back toward the product or service in order to design effective communication strategies. During the planning process, information about the customers, their wants and needs, and other background knowledge must be obtained. In IMC taxonomy, this is known as an *outside-in* approach (Kitchen, 2005; Schultz and Schultz, 2003; Hartley and Pickton, 1999; Schultz, 1996b; Stewart, 1996; Schultz, 1993e, 1993c). An important outcome of an IMC approach is that all communication mix tools, at least on the company's side, are integrated and support each other, resulting in synergy effects.

Another major driving force behind IMC is the demassification of markets. Many mass markets of the twentieth century have separated into diverse and smaller markets splintering into more specific customer segments. As a result, individual consumers' lifestyles, income classes, education, gender, and consumer actions characterize these new markets. The rich information flow from the Internet, broadcast media, and the press make information available to the consumer as never before, but research has discovered that most customers make their purchase decisions based on information of perceived value and not based only on the information a company chooses to present to them. Therefore, the ability of a company to utilize an outside-in approach and to portray a consistent image to prospects and existing customer groups becomes a critical success factor in competitive and information-rich market places (Jin, 2003/2004; Fill, 2001; Gould, Lerman, and Grein, 1999; Pickton and Hartley, 1998; Herrington and Lollar, 1996; Stewart, 1996; Schultz, Tannenbaum, and Lauterborn, 1993).

Benefits of IMC. IMC enables marketers to combine all of their communications in order to plan and create a coherent and synergistic approach. An important benefit of IMC is that it appreciates the significance to marketing strategy and corporate branding of new communication tools, such as direct marketing, Internet marketing, or different types of sales promotions (McGrath, 2005b; Pickton and Hartley, 1998; Drobis, 1997/1998).

Central to the discussion of contact points and message integration is the idea of incorporating different communication disciplines into one marketing communication campaign with the aim of achieving outcomes desirable to the company, for example, persuading customers to buy. As most organizations need to communicate with more than one target audience or stakeholder group, any campaign should take the characteristics of both the product brand and the corporate brand into consideration (Gylling and Lindberg-Repo, 2006; Phelps and Johnson, 1996; Nowak and Phelps, 1994). It is also necessary to blend every promotional mix element together as they have greater influence within an IMC approach as combined forces. Therefore, IMC can promote synergism internally among departments and, in turn, outside-in planning can be achieved (Reid, Luxton, and Mavondo, 2005; Stammerjohan *et al.*, 2005; Eagle and Kitchen, 2000; Phelps and Johnson, 1996).

Most managers perceive IMC as a process which encourages message integration and consistency, thus facilitating the interpretation of information for customers. Faced with an IMC approach, the customer will understand the different information and will not be confused by the vast amount of it from all contact points (McGrath, 2005b; Stewart, 1996; Duncan and Everett, 1993). However, if companies disregard IMC and different messages are not delivered in unison, it may lead to an incoherent brand image, which can negatively influence consumer buying and recall behavior (McGrath, 2005b; Stammerjohan *et al.*, 2005; Stewart, 1996; Schultz, 1993e). The combined consideration of product contact points and the consumer increases the company's awareness, and creates a positive relationship between the customer or prospect and the company (Gylling and Lindberg-Repo, 2006; Schultz and Kitchen, 2004; Bill, 1993). Thus, different brand-customer contact points need to be carefully maintained (McGrath, 2005b).

IMC's cost-effectiveness and its apparent ability to deliver higher return on investment (ROI) has been underlined repeatedly in the literature (Holm, 2006; Reid, 2005; Duncan and Mulhern, 2004; McGoon, 1998/1999). However, researchers have failed to explain how and why these financial benefits might be achieved.

Furthermore, other traditional marketing tools, such as the product life cycle, can be considered in the different planning stages of a product, but should not *always* be taken as a blueprint, because brands are harder to damage once customer loyalty has been achieved, for example, in the case of Apple or the smoothie brand Innocent (Kotler *et al.*, 2005). However, this is not universal, as recent product withdrawals and/or brand reputation damages have indicated. For example, the link between child labor in developing countries and global manufacturers, such as GAP, Nike, or even Coca-Cola have at least shaken consumer confidence, if not to say damaged their reputations (Thomas, 2007; Johnson and Colin, 2003). A long-term brand can only be realized if it delivers an excellent pricing, distribution, and value strategy – IMC helps in better realizing these goals (Pickton and Broderick, 2005; Cornelissen, 2001; Schultz, 1997b; Miller and Rose, 1994). IMC is valuable in that it can better inform, influence, motivate, and enlighten consumers about new and existing products because of its integrative nature and its strategic long-term focus on the actual brand (Keller, 2001).

Notably, management levels within various companies have often misinterpreted IMC, that is, as a reason to cut back on the number of company employees or to decrease the marketing budget (Low and Mohr, 1999; Schultz, 1995a). Moreover, decreasing marketing budgets, growing complexity in client-organizations, and the decreasing number of employees in organizations are conditions favorable to IMC (Pickton and Broderick, 2005; Low and Mohr, 1999). But, if a company invests in establishing an IMC approach and, for example, employs a "communication czar" as suggested by Schultz

(1991), the brands' reputation can be further strengthened and loyalty among stakeholder groups further increased. The IMC budget should be introduced as a variable cost to the finance team. Seeing it as an investment will help in understanding that if more funds are allocated to the IMC budget the outcome for the company will be greater (Reid, 2003; Low, 2000; Schultz, 1995a; Duncan and Everett, 1993).

IMC benefits can be summarized as follows (McGrath, 2005a; Reid, 2003; Pickton and Hartley, 1998):

- an IMC approach aligns short- and long-term marketing in order to avoid conflicts within an organization;
- it is a sound and clear approach;
- all target audiences are considered;
- individual and one-to-one communication is encouraged;
- synergy and recall increase;
- results in financial benefits.

Although research into IMC has discovered various benefits, their basis in reality, that is, how researchers have arrived at them, should be questioned. These benefits were either derived from quantitative studies (Eagle, Kitchen, and Bulmer, 2006; Kim, Han, and Schultz, 2004; Kallmeyer and Abratt, 2001) which left little room for an actual investigation of why these benefits have occurred, or they were the direct result of researchers credulousness when faced with marketers' positive perceptions of IMC. Most IMC research has concentrated on the opinions of advertising executives or clients. However, there has been little research directly focused on IMC benefits (Reid, 2005; Phelps and Johnson, 1996).

The Impact of Promotional Mix Elements on IMC

To develop customer-orientated messages, the manufacturing company should realize the importance of the information held by retailers and other intermediaries about customer wants and needs. One obstacle here is retailers' long-standing apprehension about ceding control to manufacturers, an attitude which has seen retailers guarding essential customer data due to fears that the manufacturer may use the information to go directly to the customer – cutting retailers and other intermediaries out of the value chain (Markillie, 2006; Reid, 2003; Schultz, 1996a). To avoid such complications between different retailers, manufacturers, and distributors, Schultz (1996a) suggested the implementation of one database where input and output access is granted to all parties. An IMC approach implies that a database should not be seen as only a storage capacity for names and addresses of customers. The advantages of a customer database include (Nowak and Phelps, 1994):

- going beyond the demographics toward the psychographics of customers;
- understanding who the most loyal and profitable customers are;
- understanding consumer attitudes, buying habits, and behavior;
- gaining more loyalty from customers.

A database can help in detecting customer preferences in relation to media channels, usage, and times. A healthy brand relationship with customers can be achieved through the use of not one promotional tool but many – such as PR campaigns, TV commercials, and magazines – in order to transmit the message (McGrath, 2005b). Multiple communication tools are not important for brand relationships alone.

On the contrary, the greater utilization of marketing communication tools within an IMC approach can lead to a better overall outcome and may also deliver a more holistic picture (Smith, Gopalakrishna, and Chatterjee, 2006; Reid, 2003; Naik and Raman, 2003; Low, 2000; Phelps and Johnson, 1996; Schultz, 1996b). Consequently, advertising alone should not be made responsible for building the brand because the brand is an asset which belongs to the whole organization (Eagle, Kithen, and Bulmer, 2006; McGrath, 2005b; Edelman, 2004; Schultz, 2004a; Drobis, 1997/1998; Hutton, 1996; Gonring, 1994; Nowak and Phelps, 1994).

Nonetheless, it is important for the marketer within the company and advertising agency to recognize the need to involve more than one promotional tool to reach different consumer groups (Economist, 2007b; Keller, 2001;

Nowak and Phelps, 1994). Ultimately, the decision-making process is controlled by the senior management and, as such, their power and involvement are essential for the development of IMC throughout the company (Reid, 2005; Pickton and Hartley, 1998). It is vital that marketers and the senior management of a company realize that an advertising agency alone will not be able to implement an IMC program because of their limited perspective and possible lack of experience with regard to coordinating and integrating all communication disciplines. Thus, the involvement of the top management must be seen as a necessity.

Therefore, the marketers of the twenty-first century should utilize an IMC program which goes beyond the usual marketing concepts, such as the 4Ps, trying, in addition, to generate new ways of creating customer segments via a database (Johnson and Schultz, 2004; Schultz and Schultz, 2003; Hutton, 1996). The informed customer of the twenty-first century will have extensive knowledge about the product category, the product per se, or even the company in general. A firm can turn the customer's search for knowledge into a competitive advantage by engaging in an interactive dialogue with the customer, for example, through direct marketing or Internet chat forums. This may then flourish into a long-term relationship in which different promotional tools can be employed to reach individual consumer groups (Cook, 2004; Johnson and Schultz, 2004; Schultz, 1995b; Nowak and Phelps, 1994).

IMC Implementation and Application

As most IMC research has been conducted with regard to advertising agencies, little is known about actual IMC application within client firms. However, it is believed that only a few organizations have reached "complete" integration (Kitchen and Li, 2005; Kitchen and Schultz, 2003; McGoon, 1998/1999). In order to draw benefits from an IMC approach, it is imperative that the whole organization, that is, across functions and subsidiaries, understands how IMC works and how IMC plans can be put into action. If this is not the case, the integration will undergo difficulties, which will in effect weaken both the corporate brand and the relationship between product brands (Reid, Luxton, and Mavondo, 2005; Davison, Bulmer, and Eagle, 2005; Duncan, 2005; Gould, Lerman, and Grein, 1999; Novelli, 1989/1990).

The next sections explore possible IMC implementation models (Pickton and Broderick, 2005; Fill, 2002; Schultz and Kitchen, 2000a). Their advantages and disadvantages are reviewed. Also reasons for embracing the Kitchen and Schultz (2001) model in this article are explained. The final section concentrates on consumer segmentation in relation to IMC, strategic planning aspects of IMC, and the kind of research that should be carried out in order for a company to act consumer-driven.

The establishment of IMC – Fill. Fill (2002) views IMC as a change of mind-set which has to be first embraced by marketers. He also points out that certain barriers exist and hence steps have to be taken to overcome such barriers. As such, the first step in an IMC approach is to focus on the promotional activities – the marketer needs to ensure that consistency and a thematic harmonization takes place among the promotional tools employed by the company (see the first box in Figure 1).

Next, the functional coordination needs to be looked after. The different parts of the organization, such as human resources, finance, corporate communications, and so on, have to be introduced to the idea of "internal marketing relationships" in order to implement an IMC approach throughout the company (see the second box in Figure 1). Hence, an IMC approach requires a cultural shift of values and beliefs, which have to come from within the organization in order for all employees to act consumer-orientated and -driven (see the third box in Figure 1). Only then can an IMC approach be fully embraced (see fourth box in Figure 1).

This model is easy to understand and also resembles certain stages of the model. However, owing to its simplicity, it lacks significant detail and explanation. It seems difficult to overcome the problem of positioning companies at their individual stages in this particular model. Furthermore, this model did not originate from actual empirical research, but rather from the critical review of current IMC literature conducted by Fill (2002).

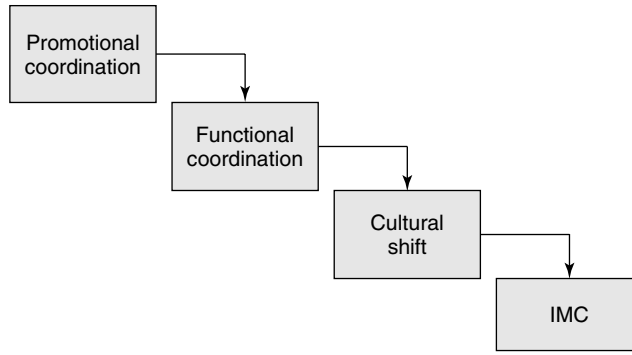


Figure 1 The establishment of IMC. (Source: Fill 2002: 469).

IMC rabostic model – Pickton and Broderick. In comparison to Fill's model (2001); Pickton and Broderick's (2005) model (see Figure 2) resulted from both a critical review of current IMC literature and empirical research (Hartley and Pickton, 1999; Pickton and Hartley, 1998).

The first step to be taken, according to Pickton and Broderick's (2005) model, is to assess the company's current position in the market place. Thus, feedback from previous marketing communication campaigns and outcomes of marketing communication activities need to be evaluated in order to determine the companies' target audience(s). The next step involves the budget allocation and the review of available resources, followed by the setting of the company's objectives and strategic plan. After determining the company's aims, decisions at the operational level have to be made. Promotional tools need to be chosen as tactical instruments, which will assist in implementing the campaign. Lastly, the success of the new campaign needs to be monitored and evaluated in order to understand if certain changes with regard to the marketing communication plan need to be made (Pickton and Broderick, 2005).

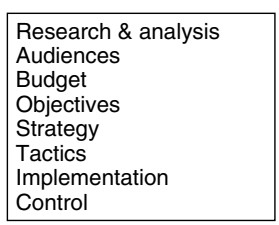


Figure 2 IMC RABOSTIC model. (Source: Pickton and Broderick 2005: 14).

Although, this model is much more detailed than that of Fill (2002), it lacks significant contribution in relation to the stages of IMC implementation. As such, the overarching aim of this model is to demonstrate how a marketing communication campaign should be implemented, and not how an IMC approach can be applied throughout an organization. In addition, not much is mentioned about the sequence of the individual stages and whether they are interchangeable. Owing to the fact that this model only concentrates on the implementation of marketing communication campaigns, it does not show how IMC could be used as an approach within an organization in the long term.

The four stages of IMC – Schultz and Kitchen. Research conducted by Kitchen and Schultz (1999) focused on advertising agencies, while best practice IMC cases were explored by the American Productivity and Quality Center (APQC) (APQC 2007) led by Schultz (Webb *et al.*, 2000; Hack *et al.*, 1999; Schultz *et al.*, 1999; Schultz and Schultz, 1998; Schultz, 1998). The findings of these studies enabled a four-stage model (see Figure 3) in order to understand the development process of an IMC approach within client organizations.

As demonstrated in Figure 3, and also by Fill (2002), the first step toward becoming integrated is to coordinate IMC at the tactical level. The aim is to generate harmony among the

various communication tools and the product. This is also known as the *one sight, one sound* approach. The primary focus is on the external communication of the brand. Most important, however, is that even in the first stage the IMC approach should be led by the company and not the agency. This is the "lowest amount" of integration a company should acquire, as it focuses on the tactical implementation of IMC. The difference between strategy and tactics is that the strategy shows which objectives the company wants to accomplish while tactics relate to how these objectives can be realized (Pickton and Broderick, 2005). If the communication between marketing communication functions, for example, sales, direct marketing, PR, and other departments (e.g., operation) fails, the customer might not understand the intended message from the company and may even perceive the company as a disappointment. However, at this level there is no mention of a customer-orientated approach.

The second level expands to a redefinition of communications, that is, all communication contact points that a customer or prospect can have with the company should be considered. Behavioral data should also be attained in order to understand customers, which is the first time that a customer-centric approach is being stressed (Kitchen and Schultz, 2001).

The third step is to utilize the increasing information flow about customer behavior, attitudes, and transactional data via a database, that is, customer data is turned into customer knowledge. Accordingly, the database should provide information about the customers, for example, how often they buy from the company or when they started buying. New technologies should also be used in order to increase the information flow between employees, distributors, and suppliers. These technologies may then help in executing messages at the right time and place (Kitchen and Schultz, 2001).

Finally, the fourth stage is to deploy IMC at the strategic level. This involves marketing and finance working together for the business to become fully customer-centric as a means of generating measurable and behavioral ROI. A successful IMC approach needs to take into consideration the corporate goals instead of simply serving tactical product brand objectives (Fitzpatrick, 2005; Schultz and Schultz, 2003; Schultz, 1997a; Gonring, 1994).

Result from previous research studies with advertising agencies (Kallmeyer and Abratt, 2001; Kitchen and Schultz, 1999; Schultz, 1997d; Miller and Rose, 1994; Caywood,

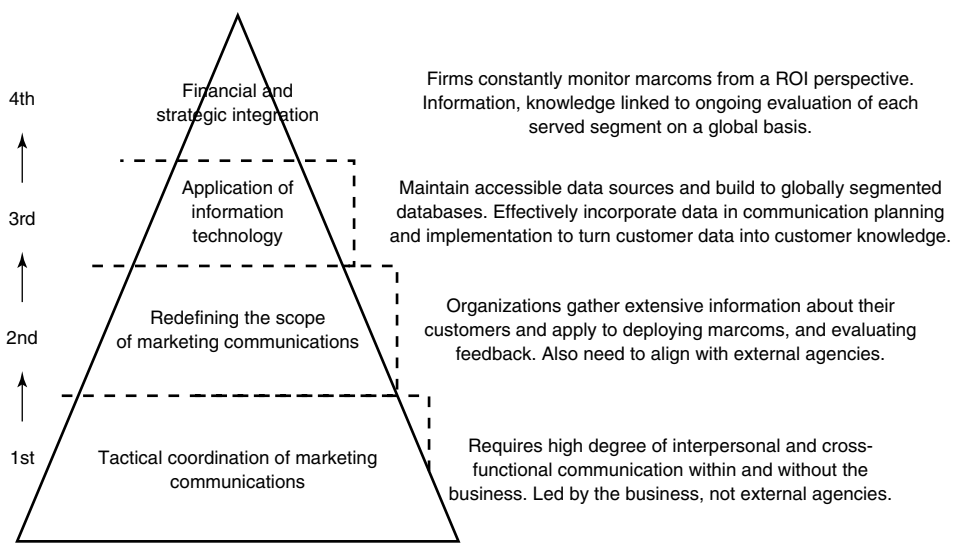


Figure 3 IMC – a four stage model. (Source: Kitchen and Schultz 2001: 108).

Schultz, and Wang, 1991b) shows that client companies may not always have the resources to implement an IMC approach at all four stages. This leads them to avoid delegating all control to a single agency, thus dividing power and responsibility among different promotional agencies. As such, during the late twentieth century, an emerging trend was reflected in the advertising agency environment with agencies forced to broaden their employees' skill sets to include a much greater variety of marketing communications skills. The aim here was to attract clients desirous of a "one-stop shopping" experience whereby agencies show a commitment to an IMC approach (Novelli, 1989/1990).

Most organizations in the early twenty-first century appear to be in stage one or two because they have yet to realize a change toward the kind of consumer-focused communication which requires close collaboration between, as a minimum, the marketing, research, and finance departments (stage four) (Kitchen *et al.*, 2004a; Schultz, 2004c). However, it should be questioned if integration among these departments is deemed to be desirable. As such, why have firms not realized this change? It may be that owing to the success of some companies, the senior management does not see it as a necessity to change current company practices.

Furthermore, it may also be possible that companies want to implement change but owing to certain reasons, for example, financial, structural or legal, the proposed change is impossible to implement. Hence, additional research is required to understand how and to what extent IMC approaches have been implemented by businesses and what kind of role agencies play as seen through the "eyes" of their clients.

In addition, financial investments for consumer research and new technologies are needed to advance through all four stages (Kitchen and Li, 2005; McGoon, 1998/1999). However, an underlying cause may also be that even though marketers and researchers have recognized the importance of IMC (Schultz, 2006b; Schultz, 2004c) and the need to focus on long-term relationship building with customers, this view has yet to be supported by the senior management. As long as IMC is perceived as a mere sales-boosting, tactical function, a company will most likely not progress through the four stages. If its strategic significance is realized, IMC can augment the corporate and product brand's image and improve relationship management between various stakeholder groups and the company (Gylling and Lindberg-Repo, 2006; Pickton and Broderick, 2005). It also helps to create a consistent marketing communication approach focusing on the long-term brand value of both corporate and product brand, which should lead to significant synergy effects. IMC, as a strategic business process, should therefore be implemented and applied throughout the organization. This view has been expressed repeatedly by these authors (Kitchen, 2005; Kitchen and Li, 2005; Kitchen *et al.*, 2004a; Schultz and Schultz, 2003; Kitchen and Schultz, 2001; 1999; Hack *et al.*, 1999; Schultz and Schultz, 1998).

After critically evaluating this model, certain points ought to be questioned. The first stage, the same as Fill's (2002) model, primarily focuses on the coordination of promotional tools in order to create a consistent marketing communication approach. However, at this stage why should the company not include customer research data in order to create tailored messages for specific target groups? Furthermore, the next three stages principally focus on the actions of the internal departments. In Kitchen and Schultz's (2001) model technologies, customer research evaluation and the coordination between agencies and the finance department are focused on, whereas in Fill's (2002) model the functional coordination and internal cultural shift are emphasized. At this point, there is no mention of the relationship the customer has with the company, be it short-term or long-term.

What is questionable about all three models is, why have the models never been adopted by any other researchers other than those who created them? Although the Kitchen and Schultz (2001) model originated from actual empirical research, it has never been tested or verified with different organizations through a wide range of industries. Thus, it is not known if this particular model is applicable to B2B and B2C industries. However, the possibility exists that because B2B companies very often have less customer numbers than B2C companies, these firms may use a much more

consumer-driven approach where databases are utilized on a day-to-day basis (Wilson, 2006; Laiderman, 2005). Hence, the model should be "tested" within actual companies in order to see, first of all, if it is applicable, and, if so, at which stage different companies find themselves to be in.

Despite such criticisms, the Kitchen and Schultz (2001) model, in comparison to the other two models, is the only model which resulted from empirical research and focuses solely on the implementation of IMCs, and not on mere marketing communications campaigns. Although Fill's (2002) model explains the different developmental stages, they appear similar to the one developed by Kitchen and Schultz (2001), only with less detail and explanation. Furthermore, Fill's model (2002) does not depict the importance of measurement in an IMC approach, which is seen as a vital component of such an approach (Alvarez, Raeside, and Jones, 2006; Reid, 2005; Kitchen *et al.*, 2004a; Eagle and Kitchen, 2000).

In summary, the previous sections have paid close attention to three different models (Pickton and Broderick, 2005; Fill, 2001; Kitchen and Schultz, 2001). A significant problem is that none of these models have been tested or validated in any form. In addition, none of these models have fostered an understanding of how IMC could be applied within companies. Albeit only partially, at least, the proposed model by Kitchen and Schultz (2001) tries to give some detail on how to implement IMC within companies. The other two models are lacking in explanations and detail, and appear at this point of time, too simplistic to be adopted as models for the empirical investigation of this thesis. Part of this thesis' empirical phase will be allocated to both reviewing the model proposed by Kitchen and Schultz (2001) and by analyzing feedback from respondents and also locating specific companies within the model.

Strategic IMC Planning

Johnson and Schulz (2004) have given some advice on how to overcome the aforementioned problems faced by organizations and agencies, and how to become a successful twenty-first century company through the application of an IMC approach. Their findings correspond with the stages outlined by Kitchen and Schultz (2001):

- conduct market research with the aim to better understand the customers;
- understand the perspectives, motivations, and behavior of each individual customer;
- focus on the customer and not product lines;
- require responsibility from the chief customer officer/chief executive officer to maximize customer value.

The four stage model (see Figure 3) reveals that an IMC plan should derive from the contact points each company has with its customers. Typically, tracking studies are used to measure all the contacts a customer has on a daily basis, for example, different media channels with a particular firm. This should give the company a general overview of who sees what kind of advertisements in the course of a given day. The marketer should then be able to divide the "mass" into segments. These segments of consumers can then be presented with targeted messages. This can be achieved with the help of sales representatives' knowledge as well as through the utilization of a fully functional database.

Recently, however, the climate of changing demographics and sociographics and rapidly expanding media choices have meant that traditional forms of segmentation, which have usually been based on consumer demographics, have been called into question. As a result, IMC researchers have identified the need for more sophisticated, behavior-orientated segmentation guidelines (Reid, 2005; Rogerson, 2005; Duncan and Mulhern, 2004; Yarbrough, 1996).

After grouping the customers into segments, promotional activities should be integrated such that they work according to the overall strategic IMC plan in order to accomplish the set marketing communication objectives. It must not be overlooked that communication between customers and organization should be a two-way process, that is, a dialogue or an exchange of information (Schultz, 2007; Smith, Gopalakrishna, and Chatterjee, 2006; Schultz, Tannenbaum, and Lauterborn, 1993). Each communication constituent may have specific aims but the end result should be

an integrated approach to the company's marketing communication activities (Pickton and Broderick, 2005).

After the objectives have (or have not) been realized, tests should be carried out in order to understand what has occurred in terms of the awareness, attitude, and behavior of customers. From such tests, areas for improvement can be identified. However, if the customers' behavior or attitude did not positively change toward the brand despite the implementation of an IMC approach, the organization should review its communication activities, such as its promotional tools, and examine its existing consumer groups. Thus, market research can again help in discovering and correcting errors (Duncan and Mulhern, 2004; Johnson and Schultz, 2004; Caywood, Schultz, and Wang, 1991b).

Nowak and Phelps (1994) developed a model (see Figure 4) which provides guidance for the correct employment of IMC at the strategic and tactical level. The authors have divided the aims of IMC into image and behavior orientation, whereas the center of the communication setting is represented by an accurate and up-to-date database which involves the market, message, and media strategy at the strategic and tactical level.

The model depicts changes in image perception and consumer behavior following a company's adoption of the IMC concept. Although advertising researchers have long recognized the merits and shortcomings of individual promotional tools, researchers (Nowak and Phelps, 1994) propose that when different promotional activities are used in an integrated manner, IMC can take place at the communication campaign (strategic level) and advertising (tactical) level. IMC, as shown in Figure 4, also encourages the employment of multiple messages within one campaign or advertisement in order to reach various target audiences (Schultz, 2007; Calder and Malthouse, 2005; Schultz, Tannenbaum, and Lauterborn, 1993) – this is also known in the early IMC literature as the *one voice approach* (Nowak and Phelps, 1994).

The significance of this model (Figure 4) in the IMC literature is apparent in its durability, having been adopted by researchers more than a decade after its emergence. For example, Grove, Carlson, and Dorsch (2007) have used it to measure IMC in advertisements at the tactical level over a 20-year time period.

Figure 4 thus demonstrates how companies can become more consumer-driven and provides examples of what activities can be used to create an integrated communication mix at the tactical advertising and strategic communication campaign level. Figure 4 aims to facilitate the implementation of an IMC approach in which individual steps show the reach of IMC implementation within a company. However, as shown in the previous sections, very little is known about IMC implementation or the strategic and tactical utilization of multiple communication tools.

In addition, although the literature on global, local, or "glocal" advertising strategies is vast, consensus with regard to the ideal option of an advertising campaign has never been reached (Wells, Burnett, and Moriarty, 2003; Kitchen and Eagle, 2002; Gould, Lerman, and Grein, 1999; Sriram and Gopalakrishna, 1991; Levitt, 1983). Thus, it is not known if an IMC approach should be adopted at a global level or if respective countries can still implement and apply their own IMC approach. As such, this research addresses the above discussed issues in the empirical phase. The next section aims to provide an overview of IMC research by reviewing some of the most profound studies in the field.

Barriers to IMC

Following up on past IMC studies, researchers were able to identify perceived barriers to IMC implementation and application. The following points represent a summary of the most important barriers to integration. These barriers are discussed in the subsequent paragraphs (highlighted in italic) (Ratnatunga and Ewing, 2005; Kitchen *et al.*, 2004a; Swain, 2004; Eagle and Kitchen, 2000; Hartley and Pickton, 1999; Pickton and Hartley, 1998; Bruhn, 1997/1998; Gonring, 1994; Duncan and Everett, 1993; Schultz, Tannenbaum, and Lauterborn, 1993):

- power, coordination, and control issues;
- client skills, centralization, and cultural issues;
- agency skills and general time/resource issues;

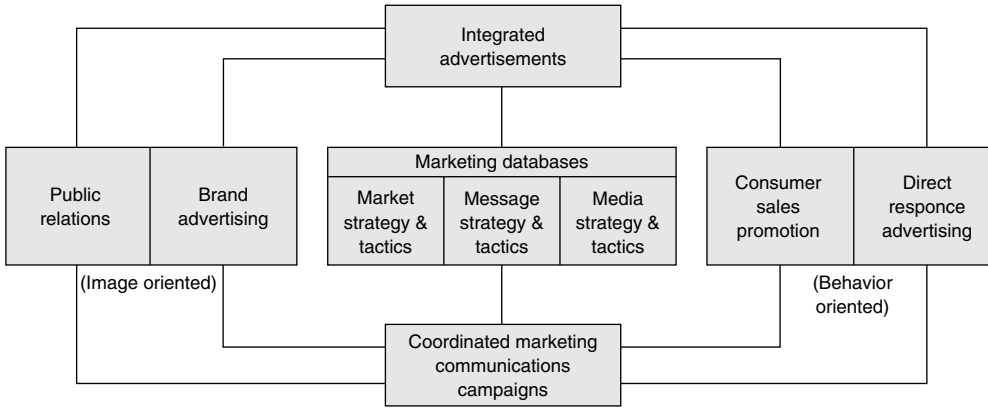


Figure 4 The integrated communications setting. (Source: Nowak and Phelps 1994: 57).

- flexibility/modification issues;
- IMC measurement difficulties;
- lack of strategic vision and business culture;
- unclear positioning at corporate level;
- subcultures within different communication departments;
- need for cross-disciplinary managerial skills;
- fear of change;
- hierarchical organization structure;
- turf battles and functional silos;
  - need to protect status;
  - need for single budgeting process and shared performance;
  - need for equal consideration across all functional areas;
  - cross-functional training;
  - inside-out communication planning;
  - short-term planning;
  - lack of database development;
- media fragmentation;
  - IMC must continually demonstrate how their mix on nontraditional media creates more impact at less cost;
- mind-set – specialization, history, tradition, experience.

IMC involves the whole company, starting with the CEO and cascading through the organization. Ideally, the corporate objectives should support the marketing objectives and vice versa. If *commitment is only given at the tactical level*, which mainly involves the marketing functions, the corporate level, such as the senior managers or the CEO itself, will probably fail to be as dedicated to the program as they should be because they will not be directly involved with the IMC implementation and may not recognize the benefits of applying an IMC approach as a strategic process (Holm, 2006; Reid, 2005; Bruhn, 1997/1998). However, even though IMC should start at the senior level, research has indicated that this occurs only very rarely (Swain, 2004). Senior managers ought to be highly engaged in the application of IMC procedures throughout the organization, and this research further explores this idea.

IMC utilizes traditional promotional elements but it is also forced to go beyond them due to the fragmentation of media and rise of new technologies. This leads to a recognition of the import of nontraditional forms of marketing communication activities, including the interactive media (such as short message services) and the Internet (e.g., through employing consumer chat rooms) (Economist, 2007b; Gonring, 1994). In particular, marketing practitioners may interpret the concept of IMC not as prerequisite of company development but rather as a one-dimensional rationalization for another agency to make an easy profit (Stammerjohan *et al.*, 2005; Wolter, 1993). Thus, owing to increasing pressure from clients and *decreasing marketing budgets*, advertising agencies may exploit the concept of IMC to include more promotional mix elements in their functions, possibly resulting in greater profits for the agency. Therefore, practitioners need to

recognize that IMC can work in favor of the organization and not just the agency. Indeed, the agency should not be seen as the final solution to the company's problems, but rather as a mixture of consultant and communication expert, working toward the company's and the agency's aims and objectives (Beard, 1997; Bill, 1993).

Another barrier is represented by the *trend toward specialization in the twenty-first century*, for example, a client-organization could employ a direct marketing agency, an advertising agency, and packaging agency at the same time. It is noticeable that many specific agencies have been developed, such as advertising, PR, and promotion agencies, but what clients really need is integration. This requires a much broader vision where aims and purposes are refocused (Naik and Raman, 2003; Schultz, 1993b). Furthermore, such specialization may also exist within the client organization, that is, the different marketing communication functions are separated from each other. Moreover, the individual tasks within one department are again divided into small fragments, which may limit the success of an IMC program (Hartley and Pickton, 1999). Therefore, division of departments may not only present a physical barrier but may also fragment the tasks of employees such that an IMC approach is thwarted. The problem is how to successfully link the different functions and departments with each other in order to start communicating and become integrated.

It is likely that most organizations' structures do not suit IMC programs. Integration necessitates not only vertical communication but also horizontal communication, a state of affairs which often causes conflicts among the employees and their functions and tasks (Gould, Lerman, and Grein, 1999; Grein and Gould, 1996; Schultz, 1993a). Communication across strategic business units (SBUs), brands and departments, as well as from brand managers to senior managers and vice versa, needs to be secured, which requires an open-minded business culture (Kim, Han, and Schultz, 2004). IMC may lead employees to a *sacrifice of power*, which could cause disputes among them because of fears arising over loss of control and authority (Eagle and Kitchen, 2000; Gonring, 1994). However, implementation and application outcomes of "full integration" have yet to be addressed in the IMC literature (Pickton and Hartley, 1998; Schultz, 1993b).

Thus, employees may fear that their status will be reduced or, worse still, their positions lost. The coordination and combination of various tasks, duties, and arrangements often represents the stiffest challenge to the implementation of an IMC program (Cornelissen and Harris, 2004; Schultz, 1993b). *Conflicts among employees* and *fragmentation of tasks* can be caused by *poor database development*, *short-term planning*, or "*inside-out*" *communication* – developing the communication strategy based on the product and then working toward the customer (Reid, 2005; Reid, Luxton, and Mavondo, 2005; Kitchen *et al*., 2004b; Gonring, 1994).

The *history or background of an organization*, as well as its *traditions and experience*, can lead to difficulties when implementing an IMC program. Firstly, its structure may often be considered satisfactory, leading to a feeling that there is no real need for a change (Schultz, 1993b). However, if this situation persists and change is not being welcomed, it may result in a loss of the company's market position. Schultz (1993d) pointed out that openness to reorganization or adaptation to new organizational structures can be the key to greater integration or an obstacle to the successful implementation of IMC, depending on the company's culture.

It has been suggested (Cornelissen, 2000; Nowak and Phelps, 1994) that to overcome barriers, cross-functional tasks, such as regular meetings or introductory training, should be developed in order to leave behind authority conflicts and create team spirit among employees. Schultz (1991) came up with the post of a communication czar who is responsible for the complete implementation and application of an IMC program within the company. McArthur and Griffin (1997) also noted that it was common practice within consumer organizations to have a single person coordinating all major communication activities. But no direct reference to Schultz's (1991) concept of a communication czar was made. In addition, Schultz (1993e) also suggested that in drastic circumstances, it may be worth restructuring the whole organization with new departments

in order to reach effective outcomes and achieve appropriate changes (Schultz, 1993e).

More than a decade has passed since the first conceptual papers on IMC have emerged, but still very little is known about how IMC can help to overcome communication barriers and how IMC is implemented in firms.

Criticisms of IMC

After the discussion of past research papers and barriers to IMC, it may be necessary to draw attention to the criticisms surrounding IMC. Firstly, Cornelissen and Lock (2000) argue that practitioners are only acknowledging and perceiving the concept of IMC as important owing to its rhetorical appeal and that there is little evidence of how an IMC approach can be implemented or how the concept is being used within organizations. There is also widespread acceptance that IMC advocates have tended to overemphasize on the language being used to describe IMC, rather than grounding IMC on the basis of empirical data (Cornelissen, 2003; Cornelissen and Lock, 2000; Schultz and Kitchen, 2000b). Indeed, the idea of integration and its nature should be viewed as "common sense" by any marketing practitioner, because marketing practitioners would almost certainly tend to be pro-integration rather than against it.

But Cornelissen and Lock (2000) are not alone in critiquing IMC. Other scholars have characterized IMC as too ambiguous, partially owing to its lack of a universally agreed-upon definition and the resulting divergence of operational IMC measures (McGrath, 2005a; Cornelissen and Lock, 2000; Stewart, 1996; Phelps and Johnson, 1996; Nowak and Phelps, 1994). To overcome this validity problem, Schultz (2004b) proposed a new and revised version of IMC, adopted by the American Marketing Association (2007). In addition, it seems that most researchers have recognized the strategic importance of IMC even if evidence in relation to the realization of the strategic employment of IMC within companies is scarce.

Furthermore, Cornelissen and Lock (2000) reproached IMC as nothing more than a managerial fad. Indeed, current IMC studies, as for example Schultz and Kitchen (2000b), have noted that IMC is still in its pre-paradigm stage and cannot, therefore, be called a theory. Despite this, in the past two decades many marketing communication investigations on IMC have proven that IMC is widely accepted among marketing and PR practitioners (Kitchen and Eagle, 2002; Reid *et al.*, 2001; Kitchen and Schultz, 1999; Kitchen and Schultz, 1998; Schultz and Kitchen, 1997; Miller and Rose, 1994; Duncan and Everett, 1993).

In addition, PR-oriented literature in the early 1990s argued against integration of marketing and PR. In an article by Miller and Rose (1994), the two authors reported on the debate at the conference for Education in Journalism and Mass Communication in 1993. Other papers and reports were also reviewed, such as that by Ehling, White, and Frunig (1992) who claimed that "the public relations function of excellent organizations exists separately from the marketing function . . . " (Miller and Rose, 1994: 13). These researchers view PR as a subject on its own, serving various stakeholder groups in a different manner than marketing does. Thus, in the eyes of these practitioners and researchers, PR should not be seen as a mere function of marketing. However, IMC was and still is perceived as reality and an important subject area to both marketing and PR practitioners, albeit for sometimes different reasons (Miller and Rose, 1994; Nakra, 1991; Niederquell, 1991).

To date, more than 100 articles and papers have been published about IMC, and although it is an emerging discipline, the fundamentals of IMC are becoming established (Patti, 2005). Schultz and Kitchen (2000b) note that the validity of a concept or theory does not only consist solely in a universally accepted definition. Although one cannot come up with a definite figure for the number of investigations needed before a specific concept can be called a theory, different concepts and frameworks are needed and ought to be interrelated and linked before an actual theory emerges (Carroll and Swatman, 2000; Walsham, 1995). As such, it is not yet known if IMC will ever emerge as a theory or if it is a concept which needs to be linked to other communication concepts to produce a different theory.

Hypothesis testing in relation to the tactical integration of marketing communication tools has begun to emerge (McGrath, 2005b;

Stammerjohan *et al.*, 2005; Jin, 2003/2004). On the other hand, different theoretical constructs and concepts are still needed to lift IMC to the theoretical level. With the apparent research focus on advertising agencies, too little research has been carried out on client organizations for there to be a demonstrable link between theoretical constructs and the experience of actual firms. As such, IMC theory building as seen from the point of view of a strategic business process has only recently begun to emerge.

In the past, researchers have argued (Cornelissen and Lock, 2000; Pickton and Hartley, 1998; Hutton, 1996; van Riel, 1995) that the historical framework and the origins of IMC can be traced back to the 1970s and 1980s, with the term *IMC* simply becoming more popular during the last decade. Caywood and Ewing (1991) also acknowledged that some ambitious practitioners of the 1950s tried to implement integrated programs and failed only because of their lack of knowledge and technology. However, it is undeniable that the nature of communication has changed dramatically during the last 20 years. Today, marketing communication practitioners must utilize these new kinds of communication, for example, the Internet, networks, value chains, direct marketing campaigns, and databases, all of which have had a significant impact on companies.

In fact, the concept of IMC has grown out of such changes (Economist, 2007b; Grove, Carlson, and Dorsch, 2007; Schultz and Kitchen, 2000b). Although practitioners may have attempted to coordinate and integrate marketing communication mix elements at the tactical level even before the earliest IMC papers were published, IMC as a strategic business process focusing on long-term brand value and customer relationship management is a product of the late twentieth century. With the aforementioned changes occurring in an increasingly global context of commercial competition, the adoption of an IMC approach throughout the company appears to be valuable, if not necessary, to sustained success.

IMC is often seen as a technique that belongs to marketing practitioners rather than to the organization, whereas PR practitioners may view it as a limiting approach given that marketing does not incorporate all functions of PR (Wightman, 1999; Miller and Rose, 1994; Wolter, 1993). Most advertising or marketing practitioners are in favor of IMC, whereas public relations practitioners are often against it. This may be due to the fact that IMC tends to be seen by PR practitioners as an attempt to bring PR under the umbrella of marketing (Miller and Rose, 1994). Wightman (1999) also suggests that many advertising agencies use IMC to incorporate PR due to decreasing client budgets. The central argument put forward by Miller and Rose (1994) is that marketing is often viewed primarily as a function which serves the customer as a stakeholder, while PR directs messages to different stakeholders.

Furthermore, Cornelissen and Harris (2004), in exploring the working relationship between marketing and PR, discovered that the majority of firms have not integrated these functions and that they rarely work in unison. Ideally, these two communication functions should be combined. Thus, the difficulties that practitioners face in assimilating these two functions may be seen as a barrier and also as a flaw insofar as PR practitioners may oppose IMC in order to prevent themselves from being subsumed by the "marketing" banner, thus denying themselves access to IMC's strategic value. Then again, neither marketing nor PR practitioners should make final conclusions because despite a vast amount of theorizing, little detailed empirical work on IMC within organizations exists to support the conceptual perspectives already advanced.

Summary and Conclusion

There is little doubt that IMC is increasingly important in the twenty-first century. Technology makes integration possible and IMC programs have already been adopted by various organizations, such as FedEx and Dell. One of IMC's aims is to create synergy among the different marketing elements in order to achieve short- and long-term returns (Madhavaram, Badrinarayana, and McDonald, 2005; Naik and Raman, 2003; Stewart, 1996). Other benefits of an IMC approach are greater consistency among the various communication messages and functions, cost savings, easier working relations between different department,s and a

better utilization of media and promotional mix elements (Pickton and Broderick, 2005).

This article has focused on the concept of IMC, past research studies on IMC, barriers to its implementation, and critiques of the concept. The investigation into past research studies has proven that although practitioners have indicated that IMC is demanded from firms, little research has been carried out in relation to client firms themselves, especially in the non-English-speaking world. Thus, the progress of IMC is not yet over. In fact, it has just begun.

Bibliography

Aaker, D.A. (2004) Leveraging the corporate brand. *California Management Review*, **46** (3), 6–18.

Alan, M. (1994) In good company. *Marketing*, March 3rd, 22–24.

Alessandri, S.W. and Alessandri, T. (2004) Promoting and protecting corporate identity: the importance of organizational and industry context. *Corporate Reputation Review*, **7** (3), 252–268.

Alvarez, J.G., Raeside, R., and Jones, W.B. (2006) The importance of analysis and planning in customer relationship marketing: verification of the need for customer intelligence and modelling. *Database Marketing and Customer Strategy Management*, **13** (3), 222–230.

American Marketing Association (2007) *Integrating the Brand Communications Process*, American Marketing Association, http: //www. marketingpower. com/ search. php? SearchFor = integrated + marketing + communication+definition&Session_ID=4b4335471 c719e98d9dbe84cd0711897&Type=0 (accessed 03 March 2007).

APQC (2007) *Official Homepage*, American Productivity Quality Center (accessed 10 May 2007).

Baker, M.J. and Balmer, J.M.T. (1997) Visual identity: trappings or substance? *European Journal of Marketing*, **31** (5/6), 366–382.

Balmer, J. (2001a) The three virtues and seven deadly sins of corporate branch management. *Journal of General Management*, **27** (1), 1–17.

Balmer, J.M.T. (2001b) Corporate identity, corporate branding and corporate marketing-seeing through the fog. *European Journal of Marketing*, **35** (3/4), 248.

Balmer, J.M.T. and Gray, E.R. (2003) Corporate brands: what are they? What of them? *European Journal of Marketing*, **37** (7/8), 972–997.

Beard, F. (1996) Integrated marketing communications: new role expectations and performance issues in the client-ad relationship. *Journal of Business Research*, **37** (3), 207–215.

Beard, F. (1997) IMC use and client-ad agency relationships. *Journal of Marketing Communications*, **3** (4), 217–230.

Bernstein, D. (1989) Advertising voices corporate void. *International Journal of Advertising*, **8** (4), 315–320.

Beverland, M. and Luxton, S. (2005) Managing integrated marketing communication (IMC) through strategic decoupling. *Journal of Advertising*, **34** (4), 103–116.

Bill, D. (1993) Future lies in integrated marketing, not communications. *Marketing News*, **27** (17), 2–11.

Bruhn, M. (1997/1998) Integrated marketing communications: a german perspective. *Journal of Integrated Communications*, 1–6. http://www.medill. northwestern. edu/ pubs/ jic/ journal/ 1997-1998. htm.

Bruhn, M. and Zimmermann, A. (1993) Integrierte kommunikationsarbeit in deutschen unternehmen – ergebnisse einer unternehmensbefragung, in *Effizientes Kommunikationsmanagment: Konzepte, Beispiele und Erfahrungen aus der integrierten Unternehmenskommunikation*, (eds M. Bruhn and D.H. Dahldoff), Schäffer-Poeschel, Stuttgart.

Calder, B.J. and Malthouse, E.C. (2005) Managing media and advertising change with integrated marketing. *Journal of Advertising Research*, **45** (4), 356–361.

Carlin, S., Womack, A., Wyckoff, T. *et al*. (1999) *Strategic and Tactical Competitive Intelligence for Sales and Marketing*, American Productivity and Quality Center, http://www.apqc.org/portal/apqc/ksn/STCIexsum.pdf?paf_gear_id=contentgearhome &paf_dm = full&pageselect=contentitem & docid= 102818 (accessed 1999).

Carlson, L., Grove, S.J., and Dorsch, M.J. (2003) Services advertising and integrated marketing communications: an empirical investigation. *Journal of Current Issues and Research in Advertising*, **25** (2), 69–82.

Carlson, L., Grove, S.J., Laczniak, R.N., and Kangun, N. (1996) Does Environmental advertising reflect integrated marketing communications?: an empirical investigation. *Journal of Business Research*, **37** (3), 225–232.

Carroll, J.M. and Swatman, P.A. (2000) Structured-case: a methodological framework for building theory in information systems research. *European Journal of Information Systems*, **9** (4), 235–242.

Caywood, C. and Ewing, R. (1991) Integrated Marketing communications: a new master's degree concept. *Public Relations Review*, **17** (3), 237–244.

Caywood, C., Schultz, D.E., and Wang, P. (1991a) *Integrated Marketing Communications: A Survey of National Goods Advertisers*, unpublished report. Medill School of Journalism, Northwestern University, Bloomington, IN, 1991.

Caywood, C., Schultz, D.E., and Wang, P. (1991b) *Integrated Marketing Communications*, Northwestern University Medill School of Journalism, Evanston, Illinois.

de Chernatony, L. (1999) Brand management through narrowing the gap between brand identity and brand reputation. *Journal of Marketing Management*, **15** (1-3), 157–179.

de Chernatony, L. and Drury, S. (2004) Identifying and sustaining service Brands' values. *Journal of Marketing Communications*, **10** (2), 73–93.

Cleland, K. (1995) A lot of talk, little action on IMC. *Business Marketing*, **80**, 1–30.

Collis, J. and Hussey, R. (2003) *Business Research – A Practical Guide for Undergraduate and Postgraduate Students*, Palgrave Macmillan, Houndmills.

Cook, W.A. (2004) Editorial: IMC's fuzzy picture: breakthrough or breakdown? *Journal of Advertising Research*, **44** (1), 1–2.

Cornelissen, J.P. (2000) Integration in communication management. *Journal of Marketing Management*, **16** (6), 597–606.

Cornelissen, J.P. (2001) Integrated marketing communications and the language of marketing development. *International Journal of Advertising*, **20** (4), 483–498.

Cornelissen, J.P. (2003) Change, continuity and progress: the concept of integrated marketing communications and marketing communications practice. *Journal of Strategic Marketing*, **11** (4), 217–234.

Cornelissen, J.P. and Harris, P. (2004) Interdependencies between marketing and public relations disciplines as correlates of communication organisation. *Journal of Marketing Management*, **20** (1/2), 237–264.

Cornelissen, J.P. and Lock, A.R. (2000) Theoretical concept or management fashion? Examining the significance of IMC. *Journal of Advertising Research*, **40** (5), 7–15.

Cornelissen, J.P., Lock, A.R., and Gardener, H. (2001) The organization of external communication disciplines: an integrative framework of dimensions and determinants. *International Journal of Advertising*, **20** (1), 67–88.

Cornelissen, J.P. and Thorpe, R. (2001) The Organisation of external communication disciplines in UK companies: a conceptual and empirical analysis of dimensions and determinants. *The Journal of Business Communications*, **38** (4), 413–438.

Davison, A., Bulmer, S., and Eagle, L. (2005) *Integrated Marketing Communication Implementation in Small New Zealand Businesses*, Massey University Department of Commerce, pp. 1–19.

Dewhirst, T. and Davis, B. (2005) Brand strategy and integrated marketing communication (IMC). *Journal of Advertising*, **34** (4), 81–92.

Dolphin, R. (1999) *The Fundamentals of Corporate Communication*, Butterworth-Heinemann, Oxford.

Drobis, D.R. (1997/1998) Integrated marketing communications redefined. *Journal of Integrated Communications*, **8**, 6–10.

Duncan, T. (2005) IMC in industry: more talk than walk. *Journal of Advertising*, **34** (4), 5–6.

Duncan, T.R. and Everett, S.E. (1993) Client perceptions of integrated marketing communications. *Journal of Advertising Research*, **32** (3), 30–39.

Duncan, T.R. and Moriarty, S. (1998) A communication-based marketing model for managing relationships. *Journal of Marketing*, **62** (2), 1–13.

Duncan, T.R. and Mulhern, F. (2004) *IMC A White Paper on the Status, Scope and Future of IMC*, Northwestern University and University of Denver.

Eagle, L. and Kitchen, P.J. (2000) IMC, brand communications, and corporate cultures client/advertising agency co-ordination and cohesion. *European Journal of Marketing*, **34** (5/6), 667–686.

Eagle, L., Kitchen, P.J., and Bulmer, S. (2007) Insights into interpreting integrated marketing communications. *European Journal of Marketing*, **41** (7-8), 956–970.

Eagle, L., Kitchen, P.J., Hyde, K. *et al.* (1999) Perceptions of integrated marketing communications among marketers and ad agency executives in New Zealand. *International Journal of Advertising*, **18** (1), 89–119.

Economist (2006a) Forbidden fruit. *The Economist*, **380** (8488), 66.

Economist (2006b) The pepsi challenge. *The Economist*, **380** (8491), 55–56.

Economist (2007a) Apple – the third act. *The Economist*, **383** (8532), 81–83.

Economist (2007b) Face value – queen of Madison Avenue. *The Economist*, **382** (8517), 76.

Edelman, R. (2004) The relationship imperative. *Journal of Integrated Communications*, 7–13. htttp://www.medill. northwestern.edu/imc/studentwork/pubs/jic/journal/current.htm.

Ehling, W., White, J., and Frunig, J.E. (1992) Public Relations and Marketing Practice in *Excellence in Public Relations and Communications Management* (ed. J. Grunig), Lawrence Earlbaum. J.E, Hillsdale, New Jersey.

Evans, J.R. and Berman, B. (1987) *Marketing*, MacMillan Publishing, New York.

Fill, C. (2001) Essentially a matter of consistency: integrated marketing communications. *The Marketing Review*, **1** (4), 409–425.

Fill, C. (2002) *Marketing Communications - Contexts, Strategies and Applications*, Financial Times Prentice Hall, Harlow.

Fitzpatrick, K.R. (2005) The legal challenge of integrated marketing communication (IMC). *Journal of Advertising*, **34** (4), 93–102.

Forman, J. and Argenti, P.A. (2005) How corporate communication influences strategy implementation, reputation and the corporate brand: an exploratory qualitative study. *Corporate Reputation Review*, **8** (3), 245–264.

Garretson, J.A. and Scot, B. (2005) The role of spokescharacters as advertisement and package cues in integrated marketing communications. *Journal of Marketing*, **69** (4), 118–132.

Gonring, M.P. (1994) Putting integrated marketing communications to work today. *Public Relations Quarterly*, **39** (3), 45–48.

Gould, S.J. (2000) The state of IMC research applications. *Journal of Advertising Research*, **40** (5), 22–23.

Gould, S.J., Lerman, D.B., and Grein, A.F. (1999) Agency perceptions and practices on global IMC. *Journal of Advertising Research*, **39** (1), 7–20.

Gray, E.R. and Balmer, J.M.T. (1998) Managing corporate image and corporate reputation. *Long Range Planning*, **31** (5), 695–702.

Gregory, J.R. and McNaughton, L. (2004) Brand logic: a business case for communications. *Journal of Advertising Research*, **44** (3), 232–236.

Grein, A.F. and Gould, E.W. (1996) Globally integrated marketing communications. *Journal of Marketing Communications*, **2** (3), 141–158.

Griffin, T., McArthur, D., Yamaki, T., and Hidalgo, P. (2000) Marketing communications: examining the work of practitioners. *International Journal of Advertising*, **18** (1), 73–94.

Griffin, G.W. and Pasadeos, Y. (1998) The impact of IMC on advertising and public relations education. *Journalism and Mass Communication Educator*, **53** (2), 4–18.

Gronstedt, A. (1996) How agencies can support integrated communications. *Journal of Business Research*, **37** (3), 201–206.

Gronstedt, A. and Thorson, E. (1996) Five approaches to organize an integrated marketing communications agency. *Journal of Advertising Research*, **36** (2), 48–59.

Grove, S.J., Carlson, L., and Dorsch, M.J. (2002) Addressing Services' intangibility through integrated marketing communication: an exploratory study. *The Journal of Services Marketing*, **16** (5), 393–411.

Grove, S.J., Carlson, L., and Dorsch, M.J. (2007) Comparing the application of integrated marketing communication (IMC) in magazine ads across product type and time. *Journal of Advertising*, **36** (1), 37–54.

Grunig, J.E. and Grunig, L.A. (1998) The relationship between public relations and marketing study. *Journal of Marketing Communications*, **4** (3), 141–162.

Gylling, C. and Lindberg-Repo, K. (2006) Investigating the Links between a corporate brand and a customer brand. *Brand Management*, **13** (4/5), 257–267.

Hack, B., Cates, C.M., Swartsfager, K., Nanda, S. *et al.* (1999) *Brand Building and Communication: Power Strategies for the 21st Century*, American Productivity and Quality Services, http://www.apqc.org/portal/apqc/ksn/STCIexsum.pdf?paf_gear_id=contentgearhome & paf_dm = full & pageselect = contentitem&docid=102818 (accessed 05 July 2006).

Hartley, B. and Pickton, D. (1999) Integrated marketing communications requires a new way of thinking. *Journal of Marketing Communications*, **5** (2), 97–106.

He, H.W. and Balmer, J.M.T. (2006) Alliance brands: building corporate brands through strategic alliances? *Brand Management*, **13** (4/5), 242–256.

Herrington, D.J. and Lollar, J.G. (1996) Comparing intensity and effectiveness of marketing communications: services vs. non-services. *Journal of Advertising Research*, **36** (6), 61–72.

Holm, O. (2006) Integrated marketing communication: from tactics to strategy. *Corporate Communications*, **11** (1), 23–43.

Hulberg, J. (2006) Integrating corporate branding and socilogical paradigms: a literature study. *Brand Management*, **14** (1/2), 60–73.

Hutton, J.G. (1996) Integrated marketing communications and the evolution of marketing thought. *Journal of Business Research*, **37** (3), 155–162.

Jin, H.S. (2003/2004) Compounding consumer interest. *Journal of Advertising*, **32** (4), 29–41.

Johnson, D. and Colin, T. (2003) *International Business: Themes and Issues in the Modern Global Economy*, Routledge, London.

Johnson, C.R. and Schultz, D.E. (2004) A focus on customers. *Marketing Management*, **13** (5), 21–26.

Kallmeyer, J. and Abratt, R. (2001) Perceptions of IMC and organisational change among agencies in South Africa. *International Journal of Advertising*, **20** (3), 361–380.

Keller, K.L. (2001) Mastering the marketing communications mix: micro and macro perspectives on integrated marketing communications programs. *Journal of Marketing Management*, **17** (7/8), 819–847.

Keller, K.L. and Aaker, D.A. (1998) The impact of corporate marketing on a company's brand extensions. *Corporate Reputation Review*, **1** (4), 356–378.

Kim, I., Han, D., and Schultz, D.E. (2004) Understanding the diffusion of integrated marketing communications. *Journal of Advertising Research*, **44** (1), 31–45.

King, S. (1991) Brand-building in the 1990s. *The Journal of Consumer Marketing*, **8** (4), 43–52.

Kitchen, P.J. (1997) Was public relations a prelude to corporate communications? *Corporate Communications*, **2** (1), 22–30.

Kitchen, P.J. (1999) Chapter 3 – the evolution of marketing and marketing communications: principles and practice, in *Marketing Communications: Principles and Practice* (ed. P.J. Kitchen), International Thomson Business Press, London, pp. 18–38.

Kitchen, P.J. (2005) New paradigm IMC – under fire. *Competitiveness Review*, **15** (1), 72–80.

Kitchen, P.J., Brignell, J., Lit, T., and Jones, G.S. (2004a) The emergence of IMC: a theoretical perspective. *Journal of Advertising Research*, **44** (1), 19–30.

Kitchen, P.J., Schultz, D.E., Kim, I. *et al*. (2004b) Will agencies ever 'Get' (or understand) IMC? *European Journal of Marketing*, **38** (11/12), 1417–1436.

Kitchen, P.J. and Eagle, L. (2002) Towards a Globalized communications strategy: perceptions from New Zealand. *Marketing Intelligence and Planning*, **20** (3), 174–184.

Kitchen, P.J. and Li, T. (2005) Perceptions of integrated marketing communications: a Chinese ad and PR agency perspective. *International Journal of Advertising*, **24** (1), 51–78.

Kitchen, P.J. and de Pelsmacker, P. (2004) *Integrated Marketing Communications: A Primer*, Routledge, London and New York.

Kitchen, P.J. and Proctor, T. (2002) Communication in postmodern integrated marketing. *Corporate Communications*, **7** (3), 114–154.

Kitchen, P.J. and Schultz, D.E. (1998) IMC – A UK Ad' agency perspective. *Journal of Marketing Management*, **14** (4/5), 465–485.

Kitchen, P.J. and Schultz, D.E. (1999) A multi-country comparison of the drive for IMC. *Journal of Advertising Research*, **39** (1), 21–39.

Kitchen, P.J. and Schultz, D.E. (2001) *Raising the Corporate Umbrella*, Palgrave, Basingstoke.

Kitchen, P.J. and Schultz, D.E. (2003) Integrated corporate and product brand communication. *Advances in Competitiveness Research*, **11** (1), 66–86.

Kliatchko, J. (2001) *Integrated Marketing Communications Theory & Practice: The Case of the Philippines*, Departmento de Empresa Informativa y Estructura de la Informacion, University of Navarre, Navarre.

Kliatchko, J. (2005) Towards a new definition of integrated marketing communications (IMC). *International Journal of Advertising*, **24** (1), 7–33.

Knox, S. and Bickerton, D. (2003) The six conventions of corporate brands. *European Journal of Marketing*, **37** (7/8), 998–1016.

Kotler, P., Wong, V., Saunders, J., and Armstrong, G. (2005) *Principles of Marketing*, Harlow Pearson Education Limited.

Laforet, S. and Saunders, J. (2005) Managing brand portfolios: how strategies have changed. *Journal of Advertising Research*, **45** (3), 314–327.

Laiderman, J. (2005) A structured approach to B2B segmentation. *Journal of Database Marketing & Customer Strategy Management*, **13** (1), 64–75.

Lawrence, L., Garber, J.R., and Dotson, M.J. (2002) A method for the selection of appropriate business-to-business integrated marketing communication mixes. *Journal of Advertising Communications*, **8** (1), 1–16.

Lee, T.J. (2002) Integration, say hello to Integrity. *Strategic Communication Management*, **6** (5), 11.

Levitt, T. (1983) The globalization of markets. *Havard Business Review*, **61** (3), 92–102.

Love, M. (2006) Cutting through the clutter at Microsoft. *Strategic Communication Management*, **10** (6), 18–19.

Low, G.S. (2000) Correlates of integrated marketing communications. *Journal of Advertising Research*, **40** (3), 27–39.

Low, G.S. and Mohr, J.J. (1999) Setting advertising and promotion budgets in multi-brand companies. *Journal of Advertising Research*, **39** (1), 67–78.

Madden, K.M. and Perry, C. (2003) How do customers of a financial services institution judge its communications? *Journal of Marketing Communications*, **9** (2), 113–127.

Madhavaram, S., Badrinarayana, V., and McDonald, R.E. (2005) Integrated marketing communication (IMC) and brand identity as critical components of brand equity strategy. *Journal of Advertising*, **34** (4), 69–80.

Markillie, P. (2006) The physical internet. *The Economist*, **379** (8483), 3–20.

Markwick, N. and Fill, C. (1997) Towards a framework for managing corporate identity. *European Journal of Marketing*, **31** (5/6), 396–409.

Martin, G., Beaumont, P., Doig, R., and Pate, J. (2005) Branding: a new performance discourse for HR? *European Management Journal*, **23** (1), 76–88.

Massie, L. and Anderson, C.L. (2003) Integrating communications: is the ideal achievable? *Corporate Communications*, **8** (4), 223–228.

McArthur, D.N. and Griffin, T. (1997) A marketing management view of integrated marketing communications. *Journal of Advertising Research*, **37** (5), 19–26.

McDonald, M.H.B., De Chernatony, L., and Harris, F. (2001) Corporate marketing and service brands – moving beyond the fast-moving consumer goods model. *Journal of Marketing*, **35** (3/4), 335–347.

McGoon, C. (1998/1999) Cutting-edge companies use integrated marketing communication. *Communication World*, **16** (1), 15–19.

McGrath, J.M. (2005a) IMC at a crossroads: a theoretical review and a conceptual framework for testing. *The Marketing Management Journal*, **15** (2), 55–66.

McGrath, J.M. (2005b) A pilot study testing aspects of the integrated marketing communications concept. *Journal of Marketing Communications*, **11** (3), 1–20.

Melewar, T.C., Karaosmanoglu, E., and Paterson, D. (2005) Corporate identity: concept, components and contribution. *Journal of General Management*, **31** (1), 59–81.

Melewar, T.C. and Walker, C. (2003) Global corporate brand building: guidelines and case studies. *Brand Management*, **11** (2), 157–170.

Miller, D.A. and Rose, P.B. (1994) Integrated communications: a look at reality instead of theory. *Public Relations Quarterly*, **39** (1), 13–16.

Muzellec, L. (2006) What is in a name change? Re-joycing corporate names to create corporate brands. *Corporate Reputation Review*, **8** (8), 305–321.

Naik, P.A. and Raman, K. (2003) Understanding the impact of synergy in multimedia communications. *Journal of Marketing Research*, **40** (4), 375–388.

Nakra, P. (1991) The changing role of public relations in marketing communications. *Public Relations Quarterly*, **36** (1), 42–45.

Niederquell, M.O. (1991) Integrating the strategic benefits of public relations into the marketing mix. *Public Relations Quarterly*, **36** (1), 23–24.

Novelli, W.D. (1989/1990) One-stop shopping: some thoughts on integrated marketing communications. *Public Relations Quarterly*, **34** (4), 7–9.

Nowak, G.J. and Phelps, J. (1994) Conceptualizing the integrated marketing communications' phenomenon: an examination of its impact on advertising practices and its implications for advertising research. *Journal of Current Issues and Research in Advertising*, **16** (1), 49–66.

Patti, C. (2005) IMC: a new discipline with an old learning approach. *Journal of Advertising*, **34** (4), 7–9.

Peltier, J.W., Schibrowsky, J.A., and Schultz, D.E. (2003) Interactive integrated marketing communication: combining the power of IMC, the new media and database marketing. *International Journal of Advertising*, **22** (1), 93–115.

Phelps, J.E., Harris, T.E., and Edward, J. (1996) Exploring decision-making approaches and responsibility for developing marketing communications strategy. *Journal of Business Research*, **37** (3), 217–223.

Phelps, J.E. and Johnson, E. (1996) Entering the quagmire: examining the 'meaning' of integrated marketing communications. *Journal of Marketing Communications*, **2** (3), 159–172.

Pickton, D. and Broderick, A. (2005) *Integrated Marketing Communications*, Prentice Hall, Harlow.

Pickton, D. and Hartley, B. (1998) Measuring integration: an assessment of the quality of IMC. *International Journal of Advertising*, **17** (4), 447–465.

Pilotta, J.J., Schultz, D.E., Drenik, G., and Rist, P. (2004) Simultaneous media usage: a critical consumer orientation to media planning. *Journal of Consumer Behaviour*, **3** (3), 285–292.

Ratnatunga, J. and Ewing, M.T. (2005) The brand capability value of integrated marketing communication (IMC). *Journal of Advertising*, **34** (4), 25–40.

Reid, M. (2003) IMC-performance relationship: further insight and evidence from the Australian marketplace. *International Journal of Advertising*, **22** (2), 227–248.

Reid, M. (2005) Performance auditing of integrated marketing communication (IMC) actions and outcome. *Journal of Advertising*, **34** (4), 41–54.

Reid, M., Johnson, T., Ratcliffe, M. *et al.* (2001) Integrated marketing communications in the Australian and New Zealand wine industry. *International Journal of Advertising*, **20** (2), 239–262.

Reid, M., Luxton, S., and Mavondo, F. (2005) The relationship between integrated marketing communication, market orientation, and brand orientation. *Journal of Advertising*, **34** (4), 11–23.

van Riel, C.B.M. (1995) *Principles of Corporate Communication*, Pearson Education, Harlow.

van Riel, C.B.M. (1997) Protecting the corporate brand by orchestrated communication. *Journal of Brand Management*, **5** (6), 409–418.

van Riel, C.B.M. and Balmer, J.M.T. (1997) Corporate identity: the concept, its measurement and management. *European Journal of Marketing*, **31** (5/6), 340–355.

van Riel, C.B.M. and Berens, G. (2003) Corporate branding: the marketing perspective, in *The Future of Marketing* (ed. P.J. Kitchen), Palgrave Macmillan, Chippenham.

Rogerson, A. (2005) Incorporate publications into your marketing strategy. *Consulting to Management*, **16** (4), 40–44.

Rose, P.B. (1996) Practitioner opinions and interests regarding integrated marketing communications in selected Latin American Countries. *Journal of Marketing Communications*, **2** (3), 125–139.

Rosenbloom, B. (2004) *Marketing Channels: A Management View*, Thomson, South Western.

Schneider, L. (1998) Agencies show that IMC can be good for bottom line. *Marketing News*, **32** (10), 11.

Schultz, D.E. (1991) Integrated marketing communications. *Journal of Promotion Management*, **1** (1), 99–104.

Schultz, D.E. (1993a) The customer and the data base are the integrating forces. *Marketing News*, **27** (24), 14.

Schultz, D.E. (1993b) How to overcome the barriers to integrations. *Marketing News*, **27** (15), 16.

Schultz, D.E. (1993c) Integration helps you plan communication from outside-in. *Marketing News*, **27** (6), 12.

Schultz, D.E. (1993d) Managers still face substantial IMC question. *Marketing News*, **27** (20), 10.

Schultz, D.E. (1993e) Maybe we should start all over with an IMC organization. *Marketing News*, **27** (22), 8.

Schultz, D.E. (1993f) We simply can't afford to go back to mass marketing. *Marketing News*, **27** (4), 20.

Schultz, D.E. (1995a) How to generate an unlimited IMC budget. *Marketing News*, **29** (12), 7.

Schultz, D.E. (1995b) Traditional advertising has role to play in IMC. *Marketing News*, **29** (18), 18–19.

Schultz, D.E. (1996a) At least in Melbourne they understand IMC. *Marketing News*, **30** (10), 8.

Schultz, D.E. (1996b) The inevitability of integrated communications. *Journal of Business Research*, **37** (3), 139–146.

Schultz, D.E. (1997a) Check out your level of integration. *Marketing News*, **31** (17), 10–12.

Schultz, D.E. (1997b) IMC in the hyper-competitive marketplace. *Marketing News*, **31** (15), 37.

Schultz, D.E. (1997c) Organize IMC programs from the outside-in. *Marketing News*, **31** (22), 6.

Schultz, D.E. (1997d) Some agencies find dip in IMC pool too cold. *Marketing News*, **31** (9), 9.

Schultz, D.E. (1998) *Expanding Knowledge Management Externally: Putting Your Knowledge to Work for Customers*, APQC, http://www.apqc.org/portal/apqc/ksn?paf_gear_id=contentgearhome&paf_dm=full&pageselect=detail\&docid=100727 (accessed 10 May 2006).

Schultz, D.E. (2003) Determine outcomes first to measure efforts. *Marketing News*, **37** (18), 7.

Schultz, D.E. (2004a) A clean brand slate. *Marketing Management*, **13** (5), 10–11.

Schultz, D.E. (2004b) IMC receives more appropriate definition. *Marketing News*, **38** (15), 8–9.

Schultz, D.E. (2004c) The marginalized brand. *Marketing Management*, **13** (6), 12–13.

Schultz, D.E. (2004d) More questions than answers. *Marketing Management*, **13** (1), 10–11.

Schultz, D.E. (2004e) Understanding total brand value. *Marketing Management*, **13** (2), 10–11.

Schultz, D.E. (2006a) Consumer control integration, not marketers. *Marketing News*, **40** (5), 7.

Schultz, D.E. (2006b) Integration's new role focuses on customers. *Marketing News*, **40** (15), 8.

Schultz, D.E. (2006c) Learning by doing. *Marketing Management*, **15** (6), 12–13.

Schultz, D.E. (2007) Focus on improving six areas on to-do list. *Marketing News*, **41** (1), 7–23.

Schultz, D.E., Carlin, S., Womack, A., Wyckoff, T. *et al*. (1999) *Strategic and Tactical Competitive Intelligence for Sales and Marketing*, American Productivity and Quality Center, http://www.apqc.org/portal/apqc/ksn/ STCIexsum.pdf ? paf_gear_id= contentgearhome & paf_dm = full & pageselect = contentitem&docid=102818 (accessed on 03 Aug. 2007).

Schultz, D.E. and Kitchen, P.J. (1997) Integrated marketing communications in US advertising agencies. *Journal of Advertising Research*, **37** (5), 7–18.

Schultz, D.E. and Kitchen, P.J. (2000a) Communicating globally, New York and Hampshire, Hampshire.

Schultz, D.E. and Kitchen, P.J. (2000b) A response to 'Theoretical Concept or Management Fashion?' *Journal of Advertising Research*, **40** (5), 17–22.

Schultz, D.E. and Kitchen, P.J. (2004) Managing the changes in corporate branding and communication: closing and re-opening the corporate umbrella. *Corporate Reputation Review*, **8** (4), 347–366.

Schultz, D.E. and Schultz, H.F. (1998) Transitioning marketing communication into the twenty-first century. *Journal of Marketing Communications*, **4** (1), 9–26.

Schultz, D.E. and Schultz, H.F. (2003) *IMC-The Next Generation*, McGraw-Hill, USA.

Schultz, D.E., Tannenbaum, S.I., and Lauterborn, R.F. (1993) *Integrated Marketing Communications*, NTC Publishing Group, Lincolnwood (Chicago).

Silverman, D. (2000) *Doing Qualitative Research – A Practical Handbook*, Sage Publications, London.

Smith, T.M., Gopalakrishna, S., and Chatterjee, R. (2006) A three-stage model of integrated marketing communications at the marketing-sales interface. *Journal of Marketing Research*, **43** (3), 564–579.

Sriram, V. and Gopalakrishna, P. (1991) Can advertising be standardized among similar countries? A cluster-based analysis. *International Journal of Advertising*, **10** (2), 137–149.

Stammerjohan, C., Wood, C.M., Chang, Y., and Thorson, E. (2005) An empirical investigation of the interaction between publicity, advertising, and previous brand attitudes and knowledge. *Journal of Advertising*, **34** (4), 55–67.

Stewart, D.W. (1996) Market-back approach to the design of integrated communications programs: a change in paradigm and a focus on determinants of success. *Journal of Business Research*, **37** (3), 147–153.

Stuart, H. and Kerr, G. (1999) Marketing communication and corporate identity: are the integrated? *Journal of Marketing Communications*, **5** (4), 169–179.

Swain, W.N. (2004) Perceptions of IMC after a decade of development: who's at the wheel, and how can we measure success? *Journal of Advertising Research*, **44** (1), 46–65.

Thomas, M. (2007) *Dispatches – Coca Cola*, Channel 4, http: // www. channel4. com/ news/ articles/ dispatches/mark + thomas + on+cocacola/1068847 (accessed 29 Nov. 2007).

Varey, R.J. (1998) Locating marketing within the corporate communication managing system. *Journal of Marketing Communications*, **4** (3), 177–190.

Walsham, G. (1995) Interpretive case studies in IS research: nature and method. *European Journal of Information Systems*, **1995** (2), 7481.

Webb, R., Hasanali, F., Lemons, D. *et al.* (2000) *Best Practice Reports*, APQC, .http://www.apqc.org/portal/apqc/ksn?paf. (accessed 03 Nov. 2004).

Webster, F.E., Malter, A.J., and Ganesan, S. (2005) The decline and dispersion of marketing competence. *MIT Sloan Management Review*, **46** (4), 35–43.

Wells, W., Burnett, J., and Moriarty, S. (2003) *Advertising Principles and Practice*, Financial Times Prentice Hall, New Jersey.

Wightman, B. (1999) Integrated communications: organization and education. *Public Relations Quarterly*, **44** (2), 18–22.

Wilson, D.R. (2006) Developing new business strategies in B2B markets by combining CRM concepts and online databases. *Competitiveness Review*, **16** (1), 38–43.

Wolter, L. (1993) Superficiality, ambiguity threaten IMC's implementation and future. *Marketing News*, **27** (19), 12–21.

Yarbrough, J.F. (1996) Putting the pieces together. *Sales and Marketing Management*, **148** (9), 68–74.

subliminal advertising

Anthony Grimes

The term *subliminal advertising* refers to marketing communication that is presented below the threshold of conscious perception. It is often related to words, symbols, and pictures that are embedded in another medium (e.g., a picture, TV program, or advertisement), and is hypothesized to exert a subconscious influence on consumer attitudes and behavior.

Although widely accepted in the domain of psychology for some time previously, the "commercialization" of subliminal perception (Broyles, 2006) occurred with the publication of a study by James M Vicary in *The Hidden Persuaders* (Packard, 1957). Vicary claimed to have increased the sales of popcorn and cola in a New Jersey cinema by projecting onto the screen subliminal instructions to consume these products. Despite the fact that he subsequently admitted the "experiment" was a hoax and that the results were invented, the notion of subliminal advertising continues to prick the public consciousness and remains a source of fascination and fear (see Nelson, 2008).

Empirical evidence for subliminal advertising effects, however, is equivocal at best (see Trappey, 1996). In support of this phenomenon, subliminal embeds have been shown to enhance hunger (e.g., the word "beef"), thirst (e.g., the word "Coke"), and sexual arousal (for a review, see Broyles, 2006). More recently, enhanced physiological states (e.g., thirst) have been found to provide the necessary conditions for the subliminal manipulation of product consumption and evaluation (Strahan, Spencer, and Zanna, 2002; Berridge and Winkielman, 2003). Such findings support the claim that subliminal advertising may have some merit in producing specific effects in the early phases of the consumer decision-making process (Cuperfain and Clarke, 1985; Theus, 1994; Dijksterhuis *et al*., 2005).

The somewhat limited supporting evidence, however, should be considered alongside continuous and consistent failures to demonstrate subliminal advertising effects (e.g., Champion and Turner, 1959; Kelly, 1979; Gable *et al*., 1987; Rosen and Singh, 1992). Furthermore, both experimental and applied replication of these effects has proven to be almost impossible (see Broyles, 2006). However, skepticism surrounding the use of subliminal advertising is primarily fueled, not by doubts over the *existence* of its influence, but by the notion that this is simply a weak analog of that created via supraliminal perception (Theus, 1994; Trappey, 1996). The effects are so small, it is argued, that they are easily subsumed by others in the consumption environment (see Broyles, 2006). Such claims lend contemporary support to the conclusion of Moore (1982) that subliminal advertising is, "an epiphenomenon, not worthy of any marketing application."

In light of such criticism, and despite continuing popular misconception, the influence of subliminal advertising has been largely discounted by academics, advertisers, and regulators alike (see Broyles, 2006; Nelson, 2008). This is not to say, however, that the potential for such influence has been entirely rejected, particularly in relation to the early (attitudinal) phases of consumer decision making and certain persuasive contexts. For example, Theus (1994) suggests that subliminal priming of positive self-image could enhance the effectiveness of public service advertising to reduce tobacco, alcohol, or drug abuse.

Although it remains difficult to make the case for a direct link between subliminal perception and consumer behavior, however, a much stronger argument has been made for other forms of nonconscious advertising effects: phenomena that do not relate to subliminal perception, but a lack of conscious awareness of the *processing* that occurs between exposure and outcome (see Chartrand, 2005; Dijksterhuis *et al*., 2005). Indeed, it should be stressed that, while the fleeting reference to Vicary's hoax stole the headlines, the concept of *subliminal advertising* is little more than a footnote in *The Hidden Persuaders* (Packard, 1957). Rather, the focus of Packard's text is the effect of nonconscious processing that occurs following *supraliminal* exposure to advertising stimuli. In this respect, he argued that advertising could influence perceptions, attitudes, emotions, and impulses without consumers' conscious awareness of the extent or nature of this influence. Such effects have since been the subject of a rich vein of

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

motivation research in both psychology and marketing over the last 40 years, during which time they have been demonstrated to be robust, replicable, and of potentially great significance to marketers, advertisers, and consumers.

See also *implicit consumer cognition; motivation research; persuasion; unconscious advertising effects*

Bibliography

Berridge, K.C. and Winkielman, P. (2003) What is an unconscious emotion? (The case for unconscious 'Liking'). *Cognition and Emotion*, **17** (2), 181–211.

Broyles, S.J. (2006) Subliminal advertising and the perpetual popularity of playing to people's paranoia. *The Journal of Consumer Affairs*, **40** (2), 392.

Champion, J.M. and Turner, W.W. (1959) An experimental investigation of subliminal perception. *Journal of Applied Psychology*, **43**, 382–384.

Chartrand, T.L. (2005) The role of conscious awareness in consumer behavior. *Journal of Consumer Psychology*, **15** (3), 203–210.

Cuperfain, R. and Clarke, T.K. (1985) A new perspective of subliminal perception. *Journal of Advertising*, **14**, 36–41.

Dijksterhuis, A., Smith, P.K., van Baaren, R.B., and Wigboldus, D.H.J. (2005) The unconscious consumer: effects of environment on consumer behavior. *Journal of Consumer Psychology*, **15**, 193–202.

Gable, M., Wilkens, H.T., Harris, L., and Feinberg, R. (1987) An evaluation of subliminally embedded sexual stimuli in graphics. *Journal of Advertising*, **16**, 26–31.

Kelly, S.J. (1979) Subliminal embeds in print advertising: a challenge to advertising ethics. *Journal of Advertising*, **8** (3), 20–24.

Moore, T.E. (1982) Subliminal advertising: what you see is what you get. *Journal of Marketing*, **46**, 38–47.

Nelson, M.R. (2008) The hidden persuaders: then and now. *Journal of Advertising*, **37** (1), 113.

Packard, V. (1957) *The Hidden Persuaders*, David McKay, New York.

Rosen, D. and Singh, S. (1992) An investigation of subliminal embed effect on multiple measures of advertising effectiveness. *Psychology and Marketing*, **9**, 157–173.

Strahan, E.J., Spencer, S.J., and Zanna, M.P. (2002) Subliminal priming and persuasion: striking while the iron is hot. *Journal of Experimental Social Psychology*, **38** (6), 556–568.

Theus, K.T. (1994) Subliminal advertising and the psychology of processing unconscious stimuli: a review of research. *Psychology and Marketing*, **11** (3), 271–290.

Trappey, C. (1996) A meta-analysis of consumer choice and subliminal advertising. *Psychology and Marketing*, **13** (5), 517–530.

sustainable marketing: collaborating with and cloning consumer 3.0

Heather Honea

In the twenty-first century, business will be responding both directly and indirectly to the issues and opportunities of a globalized market and work place, resource scarcity, climate change, and the digital networked world (Gitsham, 2008; Friedman, 2008). At the same time marketers will be tasked with creating, communicating, delivering value to and building relationships with multiple audiences (consumers, employees, associates). A key ally in navigating this complex business environment and pursuing integrated marketing communication will be the audience themselves. Marketers and consumers will become collaborators as a matter of necessity.

Collaborative Marketing: Rely on the Crowd of Consumers

The degree to which the digital networked environment connects and aggregates individuals has forced the "crowd" into the marketing equation. Search, choice, and consumption experiences have always been impacted by the social groups and networks in which consumers participate. However, the tools by which this influence is exerted are redefining all three. Group filtering and recommendation systems allow firms to deliver highly relevant and even personalized product offerings to their customers. Collaborative filtering techniques identify similarities in ratings patterns by users or in purchase behavior and are used to recommend consumption options. Amazon knows that if a particular consumer bought product A and B, when another consumer examines or purchases product A, that consumer should be informed that he might also like product B. Netflix's Cinematch system allows users to rate movies and aggregates these ratings to create recommendations for other like-minded users. Social bookmarking tools such as Delicious rank content based on user recommendations, so the crowd decides what news and information should be highlighted.

Relying on the audience to rank outcomes does not diminish the quality of recommendations or decisions. If there are diverse, independent, specialized, and local opinions in the crowd as well as a mechanism to convert private judgments into a decision, a crowd may actually be able to outperform other selection methods (Surowiecki, 2004). In fact, aggregation of information from the audience can actually achieve greater accuracy and relevance than by expert individuals. One such type of aggregation is a prediction market. A prediction market is an exchange in which participants vote on a possible outcome (e.g., product) by buying and selling shares similar to the stock market. The price of shares for a particular product is driven by the wisdom of the crowd, with a higher share price for those products the crowd believes have the greatest chance of success. Work by Ho and Chen (2007) contrasts expert opinions and consumer surveys regarding purchase intentions with prediction markets in order to identify potential mechanisms to select the most promising new product ideas and improve supply planning. They present examples from the movie, IT, and health care industry and show that prediction markets are highly accurate in these domains (Ho and Chen, 2007). Thus, the choices and decisions of the crowd may be relied on to provide critical market insight.

The value of the crowd, however, may be most striking relative to production. Firms face limits on what they can afford to produce due to the cost of pursuing certain activities in an organized way. Collaborative tools lower the costs of coordinating groups and extend the scope of production activities (Shirky, 2008). Wikipedia is an encyclopedia that does not rely on the expert-driven method of content generation but instead relies on production from the crowd. A 2005 study published in *Nature* ranked Wikipedia as comparable to *Encyclopedia Brittanica* in terms of error (2.92 mistakes per article for Britannica and 3.86 for Wikipedia) but with far more content (Giles, 2005). Apart from cost considerations, markets and audiences must be of sufficient size to justify the allocation of firm resources to production. In the past, niche or local information was not available beyond the small community who participated in or experienced an event. Photosharing sites

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

such as Flickr and blogs allow for multiple perspectives and widespread coverage of all news and niche events. With the democratization of production, where everyone has the tools to produce and create, niche products and content are made available by the crowd – dramatically increasing the population and type of goods available (Anderson, 2006).

Giving over control to the audience does come with risk; and the challenges of relying on consumers for strategy or tactics are well documented. Chevrolet's "Create your own Chevy Tahoe commercial" is one example. This campaign made available online graphic and audio assets as well as text insertion capabilities so that users could create their own Chevy Tahoe advertisements. As opposed to engaging Tahoe enthusiasts as the company expected, the most widely circulated commercials critized the SUV for its gas mileage and its negative impact on the planet (Bosman, 2006). When the crowd, or more aptly the mob, takes over content generation or messaging, the outcome is rarely as planned.

Participatory Marketing: The Integrated Consumer

The crowd is made of up an "authoritative new consumer who creates, shares, and influences" (Mooney and Rollins, 2008) and these consumers are insinuating themselves into every aspect of the marketing process. Individuals that contribute as well as consume are referred to as *prosumers* (Kelly, 2005). The *prosumer* is differentiated from the traditional consumer in that he/she is active (as opposed to passive) in terms of the business, marketing, and production activities. The prosumer may be an amateur providing professional quality services such as photography (professional consumer), the writer on Wikipedia that produces content but also consumes it (producer consumer), or the independent specialist not associated with a particular institution that provides expert support (provider consumer) (Prosumer, 2009, Retrieved September 10, 2009).

Prosumers are not just partners in the cocreation of goods and services. They have taken control of the supply side of markets. User-generated content fuels the material for video sites, participants' commentary determines the vibrancy of a blog, and the users of social and professional networking sites such as Facebook and LinkedIn determine the attractiveness of the site. In metaverses such as Second Life, participants act as architects that "build" structures, designers that sell high fashion, and artist that create works for galleries. These participants enhance the experience for other users and in some cases are compensated by those users for the goods and services they provide.

In the brick-and-mortar world, consumers managed minimal distribution activity, selling their goods at flea markets and annual garage sales. Occasionally, they placed an advertisement in the Penny Saver or classifieds but usually ended up making a trip to Goodwill. Now, entire consumer-to-consumer (C2C) markets exist, such as eBay and Craigslist. Previously, consumers relied on expert sales agents who had access to proprietary systems to facilitate their purchases. A real estate agent provided them housing listings, a travel agent detailed potential itineraries and booked their travel, a salesman duped them into a car purchase, and a doctor explained their medical condition. Today, prosumers vet houses on Realtor.com, read hotel reviews and purchase airline tickets online, research car manufacturer pricing to dealerships on Edmunds.com, and self-diagnose at WebMD.

Prosumers are able to negotiate purchase price from a position of strength. With a touch of a button they can compare product price across 10 different stores or bid at their preferred price for a product at auction. They leverage collective buying power to gain volume discounts. At Groupon.com users can take advantage of a group coupon and so long as a certain number of buyers agree to participate, they enjoy the benefits of a discounted price. Moreover, many consumers have effectively transferred transportation costs to the seller. They conveniently skip the trip to the local retailer who traditionally warehoused products and instead have items delivered directly to their door (free shipping included).

The Consumer 3.0: Narcissistic Activist

At the same time consumers take control of the marketing mix, the availability of free goods and

services (both legal and illegal) is changing their price expectations. Consumers have access to free email accounts, free entertainment content on YouTube, and free news from almost every newspaper and magazine except the Wall Street Journal. So while a recent study by the Pew Research Center shows that newspaper and news magazine readership is up, these media outlets are in crisis because more people in the United States get their news online for free than pay for it (Isaacson, 2009). Free is the new expected price point for consumers (Anderson, 2008).

Technology advances have fueled notable shifts in psychographic and behavioral characteristics of consumers as well. The mental environmental movement, a phrase coined by Kalle Lasn of Adbusters, refers to actions consumers take to eliminate mental pollution resulting from the overwhelming supply of media and marketing messages they are bombarded with daily. As the technology becomes available to filter and skip marketing communications, consumers rely on pop-up blockers, spam filters, and personal video-recorders to fast-forward through advertisements. In the best of scenarios for marketers, consumers program their own entertainment experience at Hulu (a website that offers commercial-supported streaming video of TV shows and movies from the major networks), selecting when and what they are going to watch and determining the slotting of the commercials they view.

At the same time that they adopt better tools for filtering marketing communication, consumers are perpetually connected to the digital world, multitasking during every free moment, and making themselves available for communication across larger social networks. Most striking in this dynamic is not the attention deficit or extension of social connections, but the exaggerated focus on self. The opportunity to turn the camera lens inward and provide a personal broadcast to the world is cultivating new levels of narcissism. Jean Twenge and her colleagues have documented increasing levels of narcissism in college and high school students compared to previous generations. Narcissistic traits including an inflated sense of self linked to overconfidence, materialism, lack of empathy for others, and relationship problems, are increasing in college students (Twenge *et al.*, 2008; Twenge, 2008).

Opportunities for narcissistic exposure abound. Whether it is parents sharing updates and photos of their children on a daily blog or individuals putting their photos up on Hot or Not to have the world provide ratings on their attractiveness, citizens are sharing more and more about their personal lives than ever. The opportunity to stare at one's reflection on Facebook and then share it with everyone is tempting every age group. The fastest growing segment in social networking is women over 55 (Fastest Growing Demographic on Facebook: Women Over 55, Retrieved October 08, 2009).

Interestingly, rising levels of narcissism in the United States are matched with increasing levels of consciousness toward the environment and sustainability. The number of consumers focused on health and fitness, the environment, personal development, sustainable living, and social justice is increasing annually (French and Rogers, 2006). Consumption choices and brand affiliations reflect these consumer values. Consumers carry reusable bags to the store, seek out organic foods and natural products, and purchase carbon offsets for their travel.

As consumers focus on more sustainable consumption, they have heightened expectations that firms engage in socially responsible behavior (Yoon, Gürhan-Canli, and Bozok, 2006). It appears that newly empowered consumers will take time away from gazing at their photos on Facebook to take punitive action toward firms that fail in this domain, acting as de facto regulators of corporate and social responsibility. The type of action consumers will take (civic mode) varies as a function and form of corporate social responsibility failure – the action they take may be to send a message out on Facebook but they take action nonetheless (Russell and Honea, 2009).

Clone the Consumer: Developing Articificial-Consumer-Intelligence (A-C-I)

Marketers are not entirely powerless in the partnership they develop with prosumers. As people post their opinions, ratings, and personal information across the Internet, there is more

data than ever available about these individuals. Hourly these individuals are chronicling their activities, interests, and actions – infinitely expanding the universe of content about themselves. In the *New York Times Sunday Magazine*, Clive Thompson identified services that allow people to report on the sites they are surfing (Tumblr), the places they are traveling (Dopplr), where they are located (Loopt), or what they are doing (Twitter). The data feeds from these services allow for an incredible level of ambient awareness of the person broadcasting, by anyone interested in the information. Thompson suggests that "Each little update – each individual bit of social information – is insignificant on its own, even supremely mundane. But taken together, over time, the little snippets coalesce into a surprisingly sophisticated portrait . . . like thousands of dots making a pointillist painting. This was never before possible, because in the real world, no friend would bother to call you up and detail the sandwiches she was eating." (Thompson, 2008). Firms can gain access to a simple aggregation of this type of information from users' feeds or when an individual joins a brand's page on social networking sites. Developing such digital intimacy is impractical for many firms but certainly possible. If marketers can find a way to manage the data, they can develop a profile of a consumer that is so detailed they will know the exact time of day consumers take their breaks and what they consume for an afternoon snack.

Even if firms choose not to construct these detailed profiles, prosumers will construct virtual representations of themselves in the form of avatars. According to the Gartner Inc. IT researchers, by the end of 2011, 80% of active Internet users (and Fortune 500 enterprises) will have a "second life" in a virtual world. In massively multiplayer online role-playing games (MMPORG) such as World-of-War-Craft and metaverses such as SecondLife, individuals have great flexibility in the avatars they create. Avatars may be extensions of the self and can be imbued with idealized, actual, or experimental characteristics. A person's avatar may be created with more attractive features, features intended to closely mimic physical features, or might be a different gender than the person behind the avatar. Additionally, individuals may develop avatars that have only moderate connections to the self. A *projected other* refers to an avatar imbued with idealized characteristics of another, for example, a woman designing the perfect "boyfriend." Finally, an individual might select a pure fantasy representation such as a dragon as their graphic representation. Individuals will take actions and make choices from the perspectives of the different personas engendered by their different avatars. While marketers are well versed in appealing to the idealized or the actual self, making connections with experimental, fantasy, or projected others may be quite complex. The multiplicity of representations will offer rich data regarding consumers' preferences and curiosities, but will be akin to profiling and marketing to a schizophrenic.

Citizens do not limit their data sharing to virtual worlds and social media sites. Increasingly people have begun to share data with firms so that they can "outsource" knowledge and responsibilities to the firm. They set up automatic bill pay, and do not bother to memorize phone numbers because they are perpetually connected to such information through their mobile devices. They allow word processing programs to "finalize" the spelling of a word as they write a report or reevaluate the tone of an email if the program provides them an "abrasive alert." At their favorite online stores, customers create wish or registry lists, offering up a databased list of the top products of interest to them. Every time they update their registry, consumers are providing marketing research to firms.

As individuals share more and the network is able to track more, the consumer data grows exponentially. In concert, there is a proliferation of tools that provide consumers' data about themselves. These tools expose consumers to new products, allow consumers to explore their preferences, and help them evaluate their own behavior. A visitor to Pandora enters the name of his favorite artist and instantaneously is provided a playlist or personal music channel. The tailored list comes from an extensive analysis of the musical qualities of the song that was entered and is matched to other songs with similar

attributes. A consumer simply shares a preference and is immediately provided a set of products for sampling.

Online data tools can also be used to optimize or modify consumption and behaviors. Online calculators allow for the assessment of diet goals and carbon footprints. These calculators are integrated with purchase options. Consumers can calculate the carbon impact of a trip they are going to take and purchase offsets immediately after buying an airline ticket. Personal dashboards offer the opportunity to track, quantify, and review personal consumption patterns. For instance, Google PowerMeter receives information from utility smart meters and energy management devices and provides customers with access to their home electricity consumption right on their personal iGoogle homepage. Additionally, smart home performance systems allow users to manage media, lighting, energy, and water consumption remotely.

The ideal material for research, communication, and promotion is offered up freely to marketers by consumers on a daily basis through such data collection tools. Marketers must simply help consumers organize this information in a meaningful or useful manner for both parties. To this end, customer-intelligence technology has begun to incorporate distributed computing and machine learning. Causata is working to develop a "multichannel customer interaction platform [that will] constantly update its profiles of customers, effectively learning from any purchase or query and adding that to personal information in its database. It might know that you like skiing, wine and jazz, and be cognizant of your location and calendar. Imagine a text message that includes your spouse's name in a reminder about your upcoming anniversary because you inserted the date on a wedding registry years before" (Harris, 2009). It might also include a recommendation for a romantic restaurant that carries your favorite wine and plays jazz music.

While this type of customer intelligence is characterized as predictive analysis, unique potential exists in the collection of data from consumer profiles, personal feeds, public posts, dashboards, and calculators. Kelly (1998) argues that collection of these pieces of information, in aggregate, has the potential to represent artificial intelligence. If marketers offer appropriate tools and facilitate the collection of certain types of data, the customer profiles they build will be akin to clones of those consumers. The consumer intelligence available from such clones may offer insights into behavior and choice that consumers could not share, even if they were motivated to do so.

Sustainable Marketing: Marketers as Agents of the Prosumer as Opposed to Agents of the Firm

The days of the passive couch potato are long gone. Consumers are resistant to being pitched, expect to control every step of the marketing interaction, impact the brand, and work to ensure they receive maximum value for every exchange (Hurst, 2003; Jaglois, 2003). To address this new breed of consumer, marketers must shift from acting as an agent of the firm to serving as an agent of the consumer. Such a shift demands a change in the marketing paradigm, morphing relationship marketing and fully integrating the prosumer into the marketing process. Kotler has recommended a shift to *holistic* marketing that focuses on long-term mutually satisfying relationships and coprosperity for all stakeholders (Kotler, Jain, and Maesincee, 2002). Others have labeled the model *organic* marketing because it "markets from the audience rather than to it" (Hurst, 2003), *collaborative* marketing because it entails working together with customers to create value, and *participatory* marketing because it is marketing with customers rather than at customers.

In this article, all these concepts are encapsulated under the umbrella of *sustainable marketing* (Fuller, 1999), which is described as marketing that focuses on developing long-term mutually satisfying relationships and coprosperity for all stakeholders by collaborating with the audience on the marketing process. Envisioning integrated marketing communication and brand building as the creation of a marketing ecosystem, where artificial consumer intelligence is used to sustain the organization and at the same time support the prosumer to better their consumption decisions, will require a broader perspective of the prosumer's processing of and reactions to the consumption environment.

Narrative (versus analytical) processing involves the "integrative melding of attention, feelings, and imagery" that immerses an individual into a narrative or storyline (Green, 2004) and facilitates the mental simulation of events outlined in marketing communications (Escalas, 2007). Narrative transportation enhances persuasiveness of storyline because it heightens affective intensity and realism of the experience and mitigates counter arguing (Green and Brock, 2000). An emerging body of research suggests that people's transportation into the storyline of an advertisement serves to blur the divide between the real and hypothetical and in turn facilitates persuasion Escalas (2004, 2007). Recent work by the author along with her colleagues shows that if you connect people to the marketing narrative, while transported, they go beyond the content of the narrative and build on the storyline by personalizing their own story. People not only make the leap to picture themselves in the scene but play an active role in the construction of the scene to suit their specific tastes (McFerran *et al.*, 2009). Integrated marketing communication must allow the consumer to cocreate marketing narratives. To do this, marketers will have to work in partnership with prosumers to construct a valuable and relevant narrative, inviting all facets of the consumer's different personas to participate in the development of that narrative.

Beyond co-creating narratives, marketing must develop deep insights into consumers. As marketers attempt to clone the consumer, it is important to recognize that pursuing artificial intelligence through tracking of data and machine computation requires some understanding and incorporation of "the related physical, perceptual, social, and emotional facets of intelligence" (Jenkins, 2003). Rational logic-based processes are not the sole determinants of action and decisions, but irrational and nonrational processes are major influences as well. Such reactions involve consumer responses that are spontaneous or emotional in nature. Not only do such reactions play a critical role in rational decision-making (Picard, 2003), many consumer decisions that on the surface seem to be driven by reason are better explained by emotional responses. For example, when consumers act on promotions, how they feel about a deal may be as important as the economic benefit they receive (Honea and Dahl, 2005). In fact, when the consumption context triggers certain social emotions such as guilt and pride, these emotions may actually compete with more reason-based or economic decision factors Dahl, Honea, and Manchanda (2003, 2005). When consumers experience some contradiction in their consumption reactions, they may actually engage in a sort of double-think in their decision-making. Depending on whether individuals rely on their elaborative or spontaneous response, a consumer may have entirely different evaluations of the same product or experience (Honea, Morales, and Fitzsimons, 2006). Cloning the consumer demands attention to these dualities in consumer reactions and decisions.

Such dualities may be particularly relevant as marketers try to pair sustainable marketing with sustainable consumption. Sustainable marketing is only sustainable when complemented by sustainable consumption. Sustainable consumption is consumption that is more efficient but may also involve consuming less or differently (Cohen, 2007). As such, sustainable marketing will need to help consumers organize, optimize, modify, and improve consumption behaviors. The idea of a utility company helping consumers figure out how to use less energy, instead of more energy, is a bit of paradox. Yet, it is a new marketing reality. Assisting consumers in the level, timing, and composition of demand in a way that will help *the individual* improve consumption behaviors as opposed to simply delivering on objectives of the firm will likely be the most sustainable marketing action. Thus, a key aspect of integrated marketing communication will be the reeducation of the consumers about themselves, addressing dualities in prosumers' reactions to consumption options, and facilitating the prosumer to make the best decisions even if this requires restructuring firm goals.

Bibliography

Anderson, C. (2006) *The Long Tail: Why the Future of Business is Selling Less of More*, Hyperion, New York.

Anderson, C. (2008) Why $0.00 is the Future of Business. Wired (March).

Bosman, J. (2006) Chevy Tries a Write-your-own-ad Approach, and the Potshots Fly. The New York Times (April 4).

Cohen, M.J. (2007) Consumer credit, household financial management and sustainable consumption. *International Journal of Consumer Studies*, **31** (1), 57–65.

Dahl, D.W., Honea, H., and Manchanda, R.V. (2003) The nature of self-reported guilt in consumption contexts. *Marketing Letters*, **14**, 159–171.

Dahl, D.W., Honea, H., and Manchanda, R.V. (2005) Three Rs of interpersonal consumer guilt: relationship, reciprocity, reparation. *Journal of Consumer Psychology*, **15** (4), 307–315.

Escalas, J.E. (2004) Imagine yourself in the product: mental simulation, narrative transportation, and persuasion. *Journal of Advertising*, **33** (2), 37–48.

Escalas, J.E. (2007) Self-referencing and persuasion: narrative. *Journal of Consumer Research*, **31**, 421–429.

French, S. and Rogers, G. (2006) *Understanding the Rise of Ethical Consumerism*, Natural Marketing Institute.

Friedman, T.L. (2008) *Hot, Flat and Crowded*, Farrar, Straus, and Giroux, New York.

Fuller, D.A. (1999) *Sustainable Marketing*, Sage Publications, Thousand Oaks.

Giles, J. (2005) Internet encyclopedias go head to head. *Nature*, **438**, 900–901.

Gitsham, M. (2008) Developing the global leaders of tomorrow, Ashridge, UK.

Green, M.C. (2004) Transportation into narrative worlds: the role of prior knowledge and perceived realism. *Discourse Processes*, **38** (2), 247–266.

Green, M.C. and Brock, T.C. (2000) The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, **79** (5), 701–721.

Harris, S.D. (2009) Harris: new 'consumer-intelligence' technology will compile detailed profiles. Mercury News (Oct 8).

Ho, T.-H. and Chen, K.-Y. (2007) New product blockbusters: the magic and science of prediction markets. *California Management Review*, **50**, 144–158.

Honea, H. and Dahl, D.W. (2005) The promotion affect scale: defining the affective dimensions of promotion. *Journal of Business Research*, **58** (4), 543–551.

Honea, H., Morales, A.C., and Fitzsimons, G.J. (2006) 1=2: when singular experience leads to dissociated evaluations. *Journal of Consumer Psychology*, **16** (2), 124–134.

Hurst, B.S. (2003) The audience is leading, in *The Say They Want a Revolution* (ed. M. Jaglois), iUniverse, Lincoln, pp. 10–24.

Isaacson, W. (2009) How to Save Your Newspaper Time. Time (Feb 5).

Jaglois, M. (2003) *The Say They Want a Revolution: What Marketers Need to Know as Consumer Take Control*, iUniverse, Lincoln.

Jenkins, A. (2003) Artificial intelligence in the real world. *Futures*, **35** (7), 779–786.

Kelly, K. (1998) *New Rules for the New Economu*, Penguin Books, New York.

Kelly, K. (2005) We are the Web. Wired (Aug).

Kotler, P., Jain, D.C., and Maesincee, S. (2002) *Marketing Moves: A New Approach to Profits, Growth, and Renewal*, Harvard Business School Press.

McFerran, B., Dahl, D.W., Honea, H., and Gorn, G.J. (2009) Motivational determinants of transportation into marketing narratives. working paper.

Mooney, K. and Rollins, N. (2008) *The Open Brand*, New Riders, Berkeley.

Picard, R.W. (2003) What does it mean for a computer to "have" emotions? in *Emotions in Humans and Artifacts* (eds. R.Trappl, P. Peta, and S. Payr), MIT Press.

Prosumer (2009) In Wikipedia, the free encyclopedia, from http://en.wikipedia.org/wiki/Prosumer

Russell, D.W. and Honea, H. (2009) Vice and virtue: corporate social responsibility failures and the consumer "Regulator". working paper.

Shirky, C. (2008) *Here Comes Everybody*, Penguin Press, New York.

Smith, J. Fastest growing demographic on facebook: women over 55. (Retrieved October 08, 2009) On InsideFace-book website, from http://www.insidefacebook. com/2009/02/02/ fastest-growing-demographic-on-facebook-women-over-55/#.

Surowiecki, J. (2004) *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*, Double Day, U.S.

Thompson, C. (2008) Brave New World of Digital Intimacy. The New York Times Magazine (Sep 5).

Twenge, J.M. (2008) Further evidence of an increase in narcissism among college students. *Journal of Personality*, **76**, 919–927.

Twenge, J.M., Konrath, S., Foster, J.D. *et al.*, (2008) Egos inflating over time: a cross-temporal meta-analysis of the narcissistic personality inventory. *Journal of Personality*, **76**, 875–901.

Yoon, Y., Gürhan-Canli, Z., and Bozok, B. (2006) Inferences about others on basis of corporate associations. *Journal of the Academy of Marketing Science*, **34** (2), 167–173.

historical overview of advertising promise

Zlatko Jančič

Introduction

It is frequently argued that modern advertising is no longer working. Advertising, which has helped to shape human history and especially that of the last 500 years, is now on the decline. The reasons are many, from the plethora of new media offerings to the drastic changes in the demands of consumers. This article seeks to shed some light on whether the forecasts that advertising is in decline has merit, through a historical analysis of disruptive periods in the development of the advertising profession. It is demonstrated that the current catastrophic predictions are not now being put forth for the first time but that there have been quite a few periods in the past with similar predictions. However, advertising has always found the means to remain standing. It has proven to be extremely tenacious and creative and has always been able to adapt. It appears that advertising will emerge from the current crisis, though, of course, in a new set of robes.

In this contribution, the focus is on only one aspect of the development of advertising throughout history. It concerns the different strategies or approaches the advertiser has used to argue with the offer and create a promise. In doing so, the author follows his definition of advertising, which reads as follows:

> Advertising is planned, ordered and signed creative (mass) communication, which aims to promote the exchange process between providers and consumers, giving exercisable promises.

The starting point of this analysis is any entity that triggers advertising discourse in order to stimulate interest on the side of consumers. Each period that is studied is termed an *era of advertising*, as it represents the prevailing assumptions about the best way to resolve the communication problems of the time.

The Development of Approaches to Advertising Promise

Various authors try to classify the history of advertising. Some separate the core historical periods (Arens, 1996), while others indicate only the important dates in the historical development of advertising (Egan, 2007).

Reviews of the history of advertising in one part sometimes only deal with the predominant modes of operation at that time, for example, to that part of history that covers the use of print and/or electronic mass media. One example is that of Preston (1971), who mainly deals with the historical development of media and corresponding advertising tactics. Another example of a source regarding the advertising history in the United States is *The Mirror Makers* by Fox (1997). Unfortunately, this author focuses exclusively on the development of US advertising agencies, and only on the narrow part of the history from the nineteenth century to the 1970s. Some new insights into the contemporary development of advertising using global perspective have recently been presented by Boutlis (2000), Leiss *et al*. (2005), and especially Tungate (2007).

The premise is that the existing classifications are not sufficient. These works do not reveal all of the past as well as the modern turning points relevant to understanding the development of advertising, and much less provide clues for the prediction of its possible future. This article will therefore seek to offer a more detailed periodization. Ten approaches to planning advertising copy strategy are proposed, which are named as the *eras of advertising promise*. They represent a relative sustainability through time and surpass the level of mere fads. In doing so, it is not presumed that all the details of the long and rich history of advertising have been covered. It is hoped, however, no visible era with crucial explanatory significance for advertising history has been overlooked.

The ancient era. Although it may be an attractive idea for some researchers, advertising is not, however, a typical product of capitalism. It would be closer to the truth to say that it is the product of a civilized, cultural society. The Latin root of the word *civis* means citizen,

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

a contributor to the culture (Lat. cultura) and a spiritual, intellectual, social, and artistic life (Osborne, 2007, p. 4).

Some ancient cities had at their peak several hundred thousand inhabitants. In such agglomerations, people could not always follow the principle of self-production and were therefore forced to meet their needs by entering into numerous processes of exchange or trade. The necessary condition for the success of such cities was a surplus of agricultural products, trade development, and manufacturing and service activities. A boosting engine for large cities such as Rome was also wealthy legionaries, who required a range of services in return for their money gained in combat (Osborne, 2007). These people started buying products made by unknown others (Leiss *et al.*, 2005). Urbanization, the division of labor, and the emergence of money as a medium of exchange led to the crossover from the traditional society based on social ties and barter trading to the trade based on economic exchange. Advertising was certainly one of the essential tools in this process of exchange. Advertisements in the form of wall inscriptions, signs of craftsmen, obelisks, papyruses, ceramics, and so on, have been discovered in ancient cities such as Ur and Babylon, and later, of course, in Rome, Athens, Pompeii, and Carthage. In the ruins of an Italian city, Herculaneum, archeologists found a wall with color announcements of gladiator struggles. The wall was arranged in a way similar to today's deployment of poster areas (Preston, 1971). In ancient cities, the advertisers also implemented so-called public criers (Presbey, 1929), who were sometimes accompanied by musicians. This suggests that the scale and development of advertising may have been earlier and of greater magnitude than assumed. It is also possible to conclude that ancient advertisers understood the principles of creative copywriting well, as is evidenced by the following verse of a trader from Athens:

For eyes that are shining, for cheeks like the dawn,
for beauty that lasts after girlhood has gone,
for prices in reason, the woman who knows,
will buy her cosmetics at Aesclyptöe (from Egan, 2007, p. 2).

These examples support the thesis of two advertising histories, ancient and modern, advocated by Presbey (1929), which promotes the notion of advertising development far into the precapitalist societies.

Although there are some clear advertising artifacts of ancient times, on the basis of these, it is difficult to assess the full nature of their advertising promises. We can assume that, more likely than broadcasting, narrowcasting with a mixture of informative and persuasive communication was used with a strong presence of dialogue in the processes of exchange.

With the collapse of the Roman Empire and the influx of barbaric tribes, Europe entered the Dark Ages. In the year 400, Rome had half a million inhabitants; in the year 600, only 50 000 (Osborne, 2007, p. 163). The collapse of big cities also meant the collapse of civilization and culture. During this period, there was a drastic decline in education and literacy in Europe and with this a huge decline in the presence of advertising. History, however, reports of criers, drummers, messengers, and so on, who traveled across the country disseminating news, and often advertisements, but their impact on history remains on the periphery.

Trade in the barbaric Middle Ages was limited since the majority of inhabitants of Europe were serfs to feudal lords and trade routes were at risk due to widespread banditry and piracy. It was the forming of the Hanseatic League of guilds and cities interested in developing commerce and trade in the thirteenth century that opened a new chapter in city development. Traders under arms protected themselves against pirates, banditry, and feudal lords (Presbey, 1929). Cities started to blossom by the navigable rivers and around ports all over Europe, which allowed a new boost for transport and trade. The development of medieval cities slowly led to the new rise of mass advertising. It was, on the one hand, seen in the development of the graphic signs of craftsmen, and on the other, in the stationary role of criers, who stood in front of shops, and with loud voices declared the prices and the diversity of the owners' offers. Craftsmen used their apprentices for this work, and they caused a huge amount of noise in urban markets and streets (Preston, 1971, p. 9). However, the successful functioning of trade again required

wider literacy. In this connection, the biggest shift in advertising history was brought about by Gutenberg's (circa 1439) invention of the movable printing press, which enabled mass printing in a very short time. Thus, in England, history recorded the first advertisement flyer in 1477, placed on the door of the church in Salisbury, declaring the rules of behavior for priests. On a leaflet, among other things, was also written, "Pray do not pull down this advertisement" (Egan, 2007, p. 4). During this period, the use of leaflets spread, particularly in cases where they contained information that was either too clumsy for the city "criers" or required a longer presentation of the message (Preston, 1971).

Gutenberg's invention subsequently enabled the development of printing machines for newspapers, though they did not appear until more than a 100 years later. The first newspapers started in the port cities, such as Genoa (Il Sincero – 1597), Antwerp (Niewe Tydingen – 1605), and London (The Newes – 1622), due to the intensive trade across the seas. Only three years after the appearance of the first newspaper in England, did newspapers begin to publish the first advertisements.

Information era. The new beginnings of mass advertising were facing low levels of literacy among the population, a shortage of paper and weak distribution; therefore advertising messages were detected in leaflets, flyers, and posters much earlier than those in newspapers.

Advertisements in early newspapers were of a purely informative nature, news for themselves, mostly without titles and persuasive elements. We could compare them with today's classified advertisements. The content was largely similar: news about escaped slaves, horses, war deserters, homes and properties for sale, departures and arrivals of ships, and the first manufactured goods.

Advertisers were initially hesitant, since the circulation was often too low to achieve greater coverage of the target group. Most readers were wealthy and educated men, so there were almost no advertisements, for example, offering household items (Turner, 1952). In order to promote advertising, newspapers began to introduce innovations in the design of advertising pages. Thus, in 1726, in the United States, Benjamin Franklin, editor of The Pennsylvania Gazette, began to introduce advertisement illustrations, distinguishing advertisements by an empty space, and also by imposing capital letters on titles. John Dunlop, in the newspaper Pennsylvania Pocket and General Advertiser (first issued in 1771), invented the slogan "It pays to advertise," with which he invited indecisive entrepreneurs to use this new form of commercial activity (Hornung, 1956). At that time, depending on the country, we also start to recognize the first attempts of advertisements to use distinctive persuasive elements.

Exaggeration era. With the development of manufacturing, trade, transport, literacy, and urbanity, advertisers realized the power of print mass media. Under the assumption that black on white is no doubt true, individual advertisers began to misuse the advertising message and to publish exaggerated and even false promises.

In the absence of any statutory regulation and self-regulatory mechanisms, as early as the eighteenth century, newspapers recorded an avalanche of advertisements for miracle products and/or patent medicines (Preston, 1971). These were products and devices that supposedly cured incurable diseases or returned beauty and youth to the old. "Patent medicines" usually were not patented, nor were they medicines (Fox, 1997, p. 16). They were, however, heavily and nationally promoted. Egan (2007) compares the strength of these campaigns with the postwar advertising of cigarettes, and notes how an individual category of products can permanently harm the reputation of the entire advertising profession. Another trend of this time was the widely used technique of slandering rivals. Advertisers were, without any fear of legal action, allowed to accuse rivals of lying and deceit (Turner, 1952).

As early as 1759, in the British newspaper The Idler, one of the first critics of advertising, Dr Samuel Johnson, commented on this problem and on the problem of advertising clutter as follows:

> Advertisements are now so numerous that they are negligently perused, and it is therefore become necessary to gain attention by magnificence of promises, and by eloquence sometimes sublime

and sometimes pathetic (Johnson in Turner, 1952, p. 29).

Doctor Johnson then added another, humorous, but meaningful assessment for our discussion: "The trade of advertising is now so near to perfection that it is not easy to propose any improvement" (in Turner, 1952, p. 29).

Advertising in this period was a tool with which you could promise whatever you wanted. Editors of newspapers at the time were differently oriented to the issue of puffery. Some of them declined to publish problematic advertisements, while some others, interestingly, charged three times the usual price for such an advertisement. Individual editors even occasionally warned their readers not to be ensnared by some offers in their newspaper (Turner, 1952, pp. 99–100). To control such advertisements, the first movement for truth in advertising arose followed by the first legal injunctions against misleading advertising. At the beginning of the twentieth century, the first acts of advertising self-regulation also appeared.

Era of differentiation. With specialization and differentiation of products, which offered many versions to meet the ever-diversifying demands of consumers, especially in the period between the two World Wars, new types of professionals entered in the preparation of advertisements, including designers, market researchers and, of course, copywriters. The first large advertising agencies in the United States were founded on Madison Avenue in New York. This was a time of aggressive sales orientation of companies. At about this time, Thomas J. Barratt, one of the pioneers of "modern" advertising, forecast such a development by bluntly stating that "Any fool can make soap. It takes a clever man to sell it" (in Turner, 1952, p. 89).

Differentiated products, highlighted with design advantages, began to communicate for themselves. They became status symbols, worthy of copying, and envy (e.g., Wedgwood ceramics in Britain). The style and design of products became more important than the product itself. Thus, the price or the durability of the products became of secondary importance, and into the spotlight stepped fashion, style, luxury, and the freedom to choose (Leiss *et al.*, 2005, p. 52).

Advertising played an important role here as the bringer of new values, which were in conflict with the inherited traditional patterns, and as the educator of a new, individualistic human being. Advertising took a huge step toward artistic forms, encouraged by the simultaneous rise of art, music, fashion, and the film industry.

Also new was the process of advertising research, initially based on motivational research, a branch of psychoanalysis, which was brought to the United States by a Jewish emigrant from Vienna, Ernest Dichter (1964). Motivation research was to overcome the classical statistical research methods and led to the introduction of in-depth interviews, projective techniques, thematic apperception tests (TATs) and so on, based on the belief that consumers were not rational beings and advertisers needed to understand their subconscious motives.

Twitchell (2004, pp. 65–66) notes that those advertisers, as "attention engineers" at that time, understood the problems of people well and mastered the language of sincerity and consolation. The reason for these skills could lie in the fact that many of the advertising professionals in the United States came from the higher-middle class of intellectuals of the Protestant tradition. Twitchell (2004) explains that from a historical point of view, modern branding started as the men who had studied for the ministry turned their talents to selling machine-made products. "They knew the power of promise – large promise" (Twitchell, 2004, p. 65). Twitchell (2004) then lists famous people from advertising history, like Artemas Ward, pioneering writer of slogans, who was a son of a vicar, John Wanamaker, who is known for saying that half of advertising money is thrown away, we just don't know which half, James Webb Young, who first sold Bibles door to door and later wrote a fundamental work on advertising, J. Walter Thompson, who did something similar, and later became the founder of the famous agency, and later Rosser Reeves, son of a Methodist priest, who invented the *USP* (unique selling proposition) method, which we will describe below. From this period, we can also mention one of the first modern creative copywriters, Claude Hopkins, who described his work as follows:

"I dramatize a salesman's arguments . . . Advertising must be better than ordinary argument, just as a play must be stronger than ordinary life" (Fox, 1997, p. 54).

The era of the unique selling proposition. The post-World War II period was faced with a changed world. The giant military–industrial complex made the overnight shift to production for large crowds of consumers. The production machine began to produce huge quantities of surplus, which were needed for the new markets and new consumers. Here, advertising again played a landmark role. Under the slogans of progress and a better standard of living, the Western world faced a plethora of new products and innovations for which it was necessary to educate consumers of their availability and widest possible use. Marketers needed new assumptions about the role and functions of advertising. The answer was a new approach of how best to create the advertisement promises, namely, a *unique selling proposition* or *USP*.

The author of this approach, Rosser Reeves, was the former chair of the advertising agency Ted Bates, who argued that "consumers tend to remember just one thing from the advertisement – one strong claim, or one strong concept" (Reeves, 1970, p. 34). The argument was, in short, explained as follows:

1. Each advertisement must make a proposition to the consumer. Not just words, not just product puffery, not just show window advertising. Each advertisement must say to each reader "Buy this product and you will get this specific benefit."
2. The proposition must be one that the competition either cannot, or does not, offer.
3. The proposition must be so strong that it can move the mass millions, that is, pull over new customers to your product (Reeves, 1970, pp. 47–48).

While logic might assume that a great innovation would tend to sell itself, without information about it, this will not happen. Reeves, in conjunction with this, commented on the well-known claim by Ralf Waldo Emerson, who once stated that when a man invents a better mouse trap, the world will pave the road to his door. Reeves (1970) considered that the public must first learn about the innovation and thus defined the core mission of advertising in the mass media. Reeves also commented on the popular book by Vance Packard, *The Hidden Persuaders*, where he stated that advertising has nothing to do with a hidden persuasion, but that it "works openly, in the bare and pitiless sunlight" (Reeves, 1970, p. 70).

Image era. In the 1960s, both technological development and competition increased. As we have previously seen in history, after a period of innovation a number of companies with similar "me-too products" enter the market along with producers that are supporters of "leap-frog" technology, which innovate on the basis of already-existing inventions. During this time, there were two opposite paradigms in consumer perception: the industrial and the early postindustrial. The first is derived from the desire to build social status through imitating the success of others; the other from "hippy" revolutionary rejection of the old order and finding a credo in the search for the deeper meaning of life (or at least a glimpse of it). This was the ideal time for the reform of the advertising promise.

It is at this point that the history of advertising attributes the idea of the importance of brand image to David Ogilvy. Ogilvy believed that "Every advertisement should be thought of as a contribution to the complex symbol which is the *brand image*" (author's italics) (Ogilvy, 2008, p. 128). Ogilvy believed that companies should build a brand that will stand the test of time, and that creativity was the soul of advertising. The image of the brand was considered to be a cluster of intangible properties that have no direct connection to reality. It is a kind of halo over the product or brand, which by itself may or may not be very distinctive. Thus contrary to Rosser Reeves, Ogilvy (2008) believed that consumers are not rational – what counts is the emotion; it is necessary to affect the emotional perception and not just take into account the real characteristics of the product.

The era of image in the history of advertising is associated with the so-called "creative revolution" (Fox, 1997). Although creative advertising ideas had long been present, they were dominated by the conviction of the necessity of a

"hard-selling" approach to advertising. This approach of course did not allow ambiguity in the message, such as is for example, in the print advertising the use of indirect headlines and illustrations. These are often necessary if we want the message to possess a story or even maybe a drama. The 1960s thus brought an entirely new wind to advertising, a belief that "soft-selling" messages also work. During this time, new names appeared in advertising, such as Bill Bernbach, Helmut Krone, Julian Koenig and, of course, David Ogilvy. Their advertisements for Polaroid cameras, Hathaway shirts, Chivas Regal whiskey, Volkswagen automobiles, and so on, demonstrated that even the less-known or less-popular products and brands can achieve high visibility and brand awareness. Their success thereby justified the concept of creativity as a fundamental necessity for a modern advertising promise.

Besides the slice-of-life technique in TV commercials that was extensively used and has become a kind of synonym of the image era, one further method is well worth mentioning. It is the inherent drama of Leo Burnett. Burnett believed that advertising should be based on consumer benefits with an emphasis on the dramatic element in their expression (in Belch and Belch, 2007). Every product contains its own inherent drama, the idea or the archetype that was present at its development and launch and needs to be uncovered and creatively expressed in advertising campaigns. Some of the well-known examples of Burnett's inherent drama approach were the Jolly Green Giant, Tony the Tiger, Marlboro campaign, and so on.

The era of positioning. In the 1970s, when there was not much evidence of the upcoming information revolution, early warnings were already emerging that the world was becoming overcrowded with marketing communications. Everyone was offering everything in more and more similar ways. The question that advertising strategists were facing at that time was how to stand out from the crowd and how to establish a distinctive position in the minds of consumers: the unique position. The authors, Jack Trout and Al Ries, advertising practitioners, used the term *positioning* in the series of articles through the year 1972. Encouraged by a positive feedback and successful practical results, in 1981, the authors finally published the best-selling book *Positioning: The Battle for Your Mind*.

Trout and Ries started from the assumption that advertising does not work as strongly as it is was presumed to work. A solution to the problem of a less than presumed functioning of advertising was in the specific mode of "entry" into the minds of consumers, which they called *positioning*. Trout and Ries believed that people only remember the winners or those who are the best at relevant activities or competencies. The attention of advertisers must therefore shift from mere creativity to a strategy on how to establish the leading position in the market. Since not all companies can be absolutely the leader, they must look for ways to find arguments of "partial" or a specific leadership. Trout and Ries also suggested that it is possible to play on the future position, as demonstrated in the 1970s by Avis Car Rental's famous campaign slogan "Avis is No. 2, so we try harder" (Trout and Ries, 2001).

The era of event creation. In the mid-1980s, the Western economy was facing the limits of its growth, together with additional serious warning signs of oversaturation of advertising messages. The costs to achieve the desired market outcomes became extremely high, with no guarantees about the actual success of the campaign. During this period, the belief was formed that the success of any advertising campaign should be in creating a media event. Thus, the promise gains a broader acclaim. A milestone of this era was the campaign for Apple, which, in 1984, aired the TV spot only once, during the Super Bowl. The advertisement was directed by British film director, Ridley Scott, who, in 60 seconds, presented the illusion of Big Brother from Orwell's book "1984. " The response to this advertisement was unique. Fox (1997) reported that the next day, Apple sold $6 million worth of the Macintosh. A number of TV stations later showed the advertisement, inviting Scott and other representatives of Apple as guests to their programs. The result was the perdurable effect of invaluable free publicity, which helped the Apple Corporation to save huge funds that would otherwise have been needed for purchasing media time and space, while leveraging the message and brand. The advertisement also won all the major awards

at advertising festivals that year, again adding additional free publicity.

The phenomenon of extending the advertisement to other media with the possibility of free publicity was, of course, followed by other advertisers. During this period, there was also an increase in the production of advertisements by well-known film directors (for example, Federico Fellini for Barilla and Akira Kurosawa for Suntory whiskey). With celebrated names and sometimes huge budgets, TV spots would hopefully be discussed and written about in the mass media. Of course, none of them ever managed to repeat the success of Apple.

Publicity could be achieved by other means, as well. The 1990s saw the advent of shock advertisements (Marconi, 1997) that tested the limits of ethical soundness (for instance, advertisements for Benetton, Diesel, Calvin Klein, etc.). *Shock advertisements* are the presentation of advertising messages that startle and offend viewers by violating norms and social values. The assumption was that negative publicity is better than none. Adding to this strategy was the belief by some that the media environment had changed drastically. The media were accused of losing a sense of measure as unpleasant scenes came into TV reports, entertainment, theater and movies. Many advertisers, in order to arouse attention in their advertisements, started to use swearing, hate speech, violence, antisocial behavior, and so on.

The era of a new social responsibility. The 1990s brought with them the need for reevaluation of advertising (Arens, 1996). A new paradigm in marketing and management was dawning, seeking more socially responsible practices by companies and more focus on ethics. Proponents of this paradigm believed that advertising should offer a serious contribution to improving society and the planet. In doing so, advertisers should come nearer to deontological or duty ethics as, for instance, in journalism and not merely follow the utilitarian ethical principles. Many advertisers responded positively, and in their advertisements began making their promises more relevant by highlighting social topics concerning, for example, the relationship between human beings and the environment. One example of this new perspective was the aforementioned advertisements for Benetton. However, after a certain period, Benetton started to violate these manners of political correctness and embraced a discourse of shock-advertising methods in order, some believe, to generate publicity (Boutlis, 2000). The disclosure of the possibility of such intentions was no doubt harmful to the Benetton brand.

Results of the age of social responsibility cannot yet be determined, since it has not fully established itself. However, the rise of this era can be detected again in the first years of the new millennium due to the heightened awareness of ecological threats to the planet. That we can seriously talk about the age of social responsibility in advertising can be witnessed by a global initiative to promote socially responsible advertising, organized by the IAA (International Advertising Association (2008)). An interesting trend in recent years has also been the explosion in the number of "ethical" advertising agencies. The agencies are small, and presumably deal only with the advertising of those products, services, and companies that are ethically and ecologically sound. They are often committed to encouraging people to buy fewer goods that are detrimental to the environment (Gray, 2006).

Furthermore, more advertisers in developed countries today are becoming increasingly aware that it may be wiser to promise to meet higher-level needs such as those spiritual in nature rather than promoting mere material welfare. Naish (2008, p. 218) claims that status symbols for coming generations have completely changed. The new status symbols are time, space, leisure, balance, power, and autonomy. The emerging slogan could be: "I'm having a better time than you because my personal ecology's cooler – because I know I've got enough" (Naish, 2008, p. 218). More is becoming too much and less is becoming sufficient. This recognition of voluntary simplicity, which is seen as a growing trend in some Western societies, is a good sign for the depleted environment of our planet.

The era of dialogue. Research into media habits and attitudes toward advertising today once again raises the question: "Is it really true this time?" As advertising clutter increases, so does

the number of customers who are actively or passively avoiding advertising in the conventional mass media. A lack of confidence in traditional media and advertising itself is visible among generation X, and even more so in generations Y and Z (Vollmer and Precourt, 2008). These groups have many new media at their disposal so they are freed from traditional media, and are almost immune, and are far from being passive targets. They are multitasking and are "always on" (Vollmer and Precourt, 2008). In seeking the solutions for mass advertising, advertisers rely on promises that are presented as entertainment and on the use of common values, new ideologies, a sense of belonging to the group, humor, irony, and reflexivity. Also visible are shifts to the hyperreality and media intertextuality in postmodern design strategies.

There are other changes as well. Narrowcasting such as in-store media is competing with broadcasting and many new forms of communication under the approach of integrated marketing communications (IMC) have managed to usurp the previously dominant role of mass advertising.

More and more new products and services are now launched through the Web and not through conventional mass media. News now travels in cyberspace through forums and community media in the form of tweets, blogs, online videos, and so on. Communication is diffused in the same manner as the spread of viruses. Often, all that is needed is to bring in the right people who have a reputation and a large network within their communities. Advertising experts can try to initiate positive viruses, but must be careful and unobtrusive, otherwise, these efforts can backlash. Although the euphoria of modern digital advertising may be excessive, the advertising industry will nonetheless be forced to reconsider the forms of promises that are best suited to interactive situations and that include these new media. One thing is certain: what new generations require most are the dialogue, persuasive but exercisable promises, and meaningful relationships. These relationship marketing practices are in a way the reinvention of ancient practices of strong interdependence between producers and consumers. Could all this mean that advertising has come full circle?

Conclusion

The history of advertising in the author's view negates the supposition that advertising is an exclusive product of capitalism. This assumption is highly obscuring for the study of the evolution and development of this interesting social phenomenon. The advent of mass advertising can be much more related to the development of cities, and hence the presence of intensive trade or exchange processes. It is, therefore, first of all an urban, cultural, and civilization phenomenon, which is dependent on the development of literacy in society and, of course, on the presence of a market economy. Advertising develops with the strengthening of all of these factors.

Through the examination of the individual eras of advertising, a series of approaches to the effective formulation of the advertising promise have been shown. Although it appears that some of the earlier approaches no longer work, the overview of contemporary advertising would show that almost all of those approaches can be seen as coexisting even today.

Before concluding, let us examine the limitations of the analysis. As already mentioned, the history of advertising was, for many periods, a less well-researched one. Thus, some authors present various pieces of information about the key turning points, but often do not cite reliable sources, so the analysis also cannot be completely definitive and is subject to further development and refinement.

The core direction of future research will seek answers to two important questions. Do the contemporary trends in society, associated with the virtual deurbanization through our Internet-connected homes and the feared rise of new illiteracy of younger generations due to the excessive consumption of images portend a bleak future for advertising and society as a whole? A new antiquity perhaps, or a new Middle Ages? Or can we be much more optimistic and dismiss such concerns as completely superfluous, since the development of modern advertising communication may bring something much

more stimulating? What the future will really bring, therefore, is yet to be ascertained.

Bibliography

Arens, W.F. (1996) *Contemporary Advertising*, Irwin, Chicago.

Belch, G.E. and Belch, M.A. (2007) *Advertising and Promotion. An Integrated Marketing Communications Perspective*, 7th edn, McGraw Hill/Irwin, New York.

Boutlis, P. (2000) A theory of postmodern advertising. *International Journal of Advertising*, **19** (1), 3–23.

Dichter, E. (1964) *Handbook of Consumer Motivations. The Psychology of the World of Objects*, McGraw-Hill, New York.

Egan, J. (2007) *Marketing Communications*, Thomson, London.

Fox, S. (1997) *The Mirror Makers. A History of American Advertising & Its Creators*, Illini Books Edition, Urbana and Chicago.

Gray, R. (2006) Rise of the ethical agency. *Campaign*, **14** (7), 33.

Hornung, C.P. (1956) *Handbook of Early Advertising Art*, Dover Publications, Inc., New York.

International Advertising Association (2008) IAA Launches "Socially Responsible" Ad Award. http://www. iaauae. org/ en/ press-releases/iaa-launches-socially-repsonsible-ad-award.html (accessed on 30 May 2008).

Leiss, W., Kline, S., Jhally, S., and Botterill, J. (2005) *Social Communication in Advertising. Consumption in the Mediated Marketplace*, Routledge, London.

Marconi, J. (1997) *Shock Marketing: Advertising Influence on Family Values*, Bonus Books, Chicago.

Naish, J. (2008) *Enough. Breaking Free from the World of More*, Hoodder & Stoughton, London.

Ogilvy, D. (2008) *Confessions of an Advertising Man*, Southbank Publishing, London.

Osborne, R. (2007) *Civilization. A New History of Western World*, Pimlico, London.

Presbey, F. (1929) *The History and Development of Advertising*, Doubleday, Doran & Company, New York.

Preston, G. (1971) *Advertising*, BT Batsford Ltd, London.

Reeves, R. (1970) *Reality in Advertising*, Alfred A. Knopf, New York.

Tungate, M. (2007) *Adland. A Global History of Advertising*, Kogan Page, London.

Trout, J. and Ries, A. (2001) *Positioning. The Battle for Your Mind*, McGraw Hill, New York.

Turner, E.S. (1952) *A Shocking History of Advertising*, Michael Joseph, London.

Twitchell, J.B. (2004) *Branded Nation. The Megachurch, College Inc., and Museumworld*, Simon & Schuster, New York.

Vollmer, C. and Precourt, G. (2008) *Always on. Advertising, Marketing and Media in an Era of Consumer Control*, McGraw Hill, New York.

unconscious advertising effects

Marjolein Moorman

Most traditional advertising effect models, like AIDA and DAGMAR, are based on the premise that advertising is attended to and processed consciously. However, recent neuroscientific research shows that only a small portion of the enormous amount of information that people are confronted with each day is processed consciously. Most information is unconsciously attended to, processed, and stored in memory (Lamme, 2004; Ledoux, 1998).

In the middle of last century, scientists became aware that people are unconsciously scanning their environment for relevant information continuously. Only when we encounter relevant information do we shift from unconscious to conscious attention. This shift is suitably illustrated by a well-known social psychological phenomenon called *the cocktail party effect* (Moray, 1959). The cocktail party effect takes place in information-dense surroundings, for example, a crowded party. Despite the clamor of many voices, most people are able to concentrate on one conversation at a time. It seems that they can mentally block out all the other noise. This is until their name is mentioned on the other side of the room. Most people will notice this. Apparently they are able to consciously concentrate on the conversation while unconsciously scanning the environment for other relevant information at the same time. The moment they hear their name being mentioned at the other side of the room they are likely to shift their conscious attention toward that conversation to hear what is said about them.

Starting as an evolutionary survival mechanism, unconscious processing is a very valuable human trait these days. It helps us filter relevant information out of the daily cacophony of messages calling for our attention. Where we can concentrate consciously on only one task at a time, we are able to process other information unconsciously with a seemingly unlimited capacity (Dijksterhuis, 2007). Nøretranders (1998) has estimated that of the 11 million parts of information that the human mind processes each second, only 40 to 60 are processed consciously.

The concept of unconscious processing is very interesting for an advertising landscape that is challenged by a growing amount of messages being sent and the declining consumer attention rates. Heath (2001) poses that consumers almost never pay full attention toward advertising messages these days, because of lack of interest and a limited mental capacity to consciously process information. However, in his opinion, this does not mean that advertising is no longer effective. In fact, he poses that advertising is most effective when it is attended to and processed only on a subconscious level. The unconsciously processed brand information will have an effect on subsequent consumer attitudes and behavior, but without us being aware of it. And since we are not consciously aware, we cannot actively defend ourselves against it.

To date, the number of studies providing evidence for Heath's hypothesis that unconsciously processed advertising messages are more effective than consciously processed advertising message, remains limited. (DuPlessis, 2005, p. 141) calls this Heath's error. The problem with studying the effects of unconsciously processed advertising messages is that people are not able to remember them and that well-known advertising effect measures such as recall and recognition are thus not suitable. Physiological measures, such as fMRI and EEG scans, and implicit measures, such as implicit cognition and attitude tests, where there is no explicit mentioning of the advertisement are better suited. In the last decade, a growing number of scholars have applied these kinds of measures to study the unconscious effects of advertising. Results vary, which might be the consequence of the tests not being refined well enough to effectively study these kinds of effects. Further developments and validity tests of physiological and implicit measures are an absolute necessity to bring clarity to our understanding of unconscious advertising effects.

See also *advertising effectiveness; subliminal advertising*

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Bibliography

Dijksterhuis, A. (2007) *Het Slimme Onbewuste: Denken met Gevoel [The smart unconscious: Thinking with feelings.]*, Bert Bakker, Amsterdam.

DuPlessis, E. (2005) *Advertised Mind: Groundbreaking Insights into how Our Brains Respond to Advertising*, Kogan Page Ltd.

Heath, R. (2001) *The Hidden Power of Advertising. How Low Involvement Processing Influences the Way we Choose Brands*, Admap Publications, Oxfordshire.

Lamme, V.A.F. (2004) Separate neural definitions of visual consciousness and visual attention: a case for phenomenal awareness. *Neural Networks*, **17**, 861–872.

Ledoux, J. (1998) *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*, Weidenfeld & Nicolson, London.

Moray, N. (1959) Attention in dichotic listening: affective cues and the influence of instruction. *Quarterly Journal of Experimental Psychology*, **11**, 56–60.

Nøretranders, T. (1998) *The User Illusion: Cutting Consciousness Down to Size*, Viking, New York.

copy test methods to pretest advertisements

Cornelia Pechmann and Craig Andrews

In this paper, we define advertising copy testing, describe the major copy test methods using case examples, discuss each method's strengths and limitations, and identify areas of debate. The case examples were selected based on the authors' direct experiences and the methodological rigor of the methods. We define advertising copy testing as a quantitative and qualitative marketing research method that is used to pretest advertisements where numerical data on effects is collected and statistically analyzed. The key industry document about it called PACT (1982) states: "Copy testing is undertaken when a decision is to be made about whether the advertising should be run in the marketplace. Whether this stage utilizes a single test or a combination of tests, its purpose is to aid in the judgment of specific advertising executions" (page 8).

Generally, there are four stages in advertising research: (i) copy development, (ii) a rough stage, (iii) copy testing (i.e., pretesting), and (iv) tracking (i.e., posttesting) (Belch and Belch, 2007; Shimp, 2010). In copy development, initial ideas are developed from the creative brief and screened in focus groups. In the rough stage, the advertisement is tested in draft or storyboard form, again in focus groups. Next, the final or nearly final advertisements are copy tested. Finally, after appearing in the media, the advertisements are tracked using surveys and/or sales data. The surveys typically assess advertising recall and recognition, although some measure advertising attitudes, product beliefs, and/or purchase intent.

Nine principles of copy testing were developed by US advertising agencies and are set forth in PACT (1982, p. 10–27). PACT states that for a copy test method to be sound, it must (i) be relevant to the objectives, (ii) have agreement about the use of results, (iii) use multiple measurements, (iv) be based on a model of human response, (v) consider multiple exposures, (vi) test comparably finished executions, (vii) control the exposure context, (viii) define the relevant sample, and (ix) demonstrate reliability and validity. To this list, we add (x) take baseline (i.e., pre-exposure) measurements and/or use control groups. We add this because it is in accordance with sound and generally accepted research standards (Cook and Campbell, 1979). A copy test may use a prepost design that compares participants' responses before and after they see the advertisement. Or, a copy test may randomly assign participants to a test group that sees the advertisement, or to a control group that does not, and compare their responses. Some designs make use of both approaches (Malholtra, 2009). We now describe three major copy test methods that apply these principles that are used by the US National Youth Anti-Drug Media Campaign, the US Federal Trade Commission, 1983, and the ARSgroup, a major copy test firm.

Method 1

The US Office of National Drug Control Policy, with the assistance of The Partnership for a Drug-Free America, has run the National Youth Anti-Drug Media Campaign for about a decade. Total expenditures are over a billion dollars (Foley and Pechmann, 2004). The antidrug advertisements target either the youth or their parents. Each advertisement is copy tested prior to airing using a method that was developed by academic experts who conduct similar copy tests for their own research (e.g., Pechmann and Reibling, 2006).

For each ad tested in a youth copy test, the sample consists of 200 Caucasians, 200 Hispanics, and 200 African-Americans, split 50/50 on grade (grades 7–8 vs. grades 9–10) and 50/50 on gender. From each ethnic group and grade, 50 youths are randomly assigned to the test group (they see the antidrug advertisement) and 50 are randomly assigned to the nonexposure control group (they do not see the advertisement). For each ad tested in a parent copy test, the sample consists of 100 Caucasians, 100 Hispanics, and 100 African-Americans, with at least one child in grades 7–10, and split about 50/50 on gender. From each ethnic group, 50 parents are randomly assigned to the test group and 50 are randomly assigned to the control group. The sample sizes were chosen to detect differences of 10% or more between test and control groups with a power of 0.80, a 95% confidence level ($p < 0.05$), and an observed within-group variance of 0.6.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

The data collection is subcontracted to an experienced marketing research firm and is typically completed in one weekend using about 40 shopping malls across the United States where the firm has access to research facilities. Youth are prescreened based on grade, ethnicity, and gender and their sensation seeking is also assessed. Parents are prescreened based on their children's grades and their own ethnicity. Those who pass the screen and give informed consent are paid $1 for completing the 10-minute study and the response rate is about 78%. Parental consent for youths' participation is obtained whenever feasible. The study is conducted in small rooms with trained English-speaking interviewers who read the questions to individual participants and record their answers. Typically multi-item scales with 5 to 10 point response options are used.

Test participants see the antidrug advertisement two times in succession, while control participants do not see it. Next, test participants answer open-ended questions about what the advertisement is about and their responses are later coded. Test participants also rate the advertisement on diagnostic measures of effectiveness, believability, attention-getting power, and argument strength. The advertisement's ratings are compared with norms from prior testing. Then, all participants complete a questionnaire that includes measures of their drug-related beliefs and intent. The measures are pretested to ensure clarity and sensitivity (no ceiling/floor effects) and are reused across multiple copy tests to ensure comparability. The data are analyzed using analysis of covariance; there is one independent variable (test vs. control condition), moderators or contingent factors (for youth: grade, gender, and ethnicity; for parents: ethnicity and gender) and several covariates (for youth: exact age and sensation seeking score; for parents: the youth's grade). The effects of test versus control condition alone and in combination with grade, gender, and/or ethnicity are examined by looking at main effects and two- and three-way interactions. This analysis is done for each belief and intent measure.

A youth advertisement is recommended for airing if it significantly ($p < 0.05$) strengthens an antidrug belief or weakens intent to use drugs among the overall sample or among a subsample based on grade, gender, or ethnicity. It is not recommended for airing if it weakens an antidrug belief or strengthens intent to use drugs. A parent advertisement is recommended for airing if it significantly ($p < 0.05$) strengthens an antidrug parenting belief or intent among the overall sample or a subsample based on gender or ethnicity. It is not recommended for airing if it weakens an antidrug parenting belief or intent. The diagnostic measures help reveal an advertisement's specific strengths or weaknesses.

Strengths

The National Youth Anti-Drug Media Campaign's copy test method seems to comply with all 10 principles of good copy testing that were outlined above.

1. *Be relevant to the objectives*: The campaign objectives are to strengthen antidrug beliefs and intent and avoid adverse effects, and the copy testing explicitly tests this.
2. *Have agreement about the use of results*: There are clear criteria for recommending or not recommending advertisements for airing.
3. *Uses multiple measurements*: There are multiple measures of the ad, beliefs, and intent.
4. *Be based on a model of human response*: Well-established models are used, such as hierarchy of effects, the elaboration likelihood model, attitude toward the ad, and steps in information processing (Vakratsas and Ambler, 1999). For example, beliefs are predicted to affect intent.
5. *Consider multiple exposures*: Participants see the antidrug advertisement twice.
6. *Test comparably finished executions*: Virtually all advertisements are tested in final form.
7. *Control the exposure context*: The antidrug advertisement is shown alone, without clutter.
8. *Define the relevant sample*: The target grades, genders, and ethnicities are stated.
9. *Demonstrate reliability and validity*: Multi-item measures are used to assess reliability. Also, copy test results are compared with tracking results to assess validity. Both reliability and validity have been found to be high.

10. *Take baseline (i.e., pre-exposure) measurements and/or use control groups*: A randomly selected control group does not see the antidrug advertisement, but just completes the belief and intent measures, and statistical comparisons are made between control and test groups. Compared with prepost designs, control groups help rule out alternative explanations for reported antidrug beliefs and intent, such as mere measurement effects and cueing of other outside campaigns simply caused by asking about antidrug beliefs and intent. In addition, a prepost design was ruled out because pretest responses are likely to bias posttest responses; the measurements are too close in time (Cook and Campbell, 1979). Specific biases of concern are that people may purposely report the same beliefs at both points in time or exhibit reactance against the testing and claim adverse effects.

Limitations

In the National Youth Anti-Drug Media Campaign's copy tests, youths' interest in drug use is not measured. Sensation seeking, which is related to drug use, is measured based on youth activities and interests (e.g., "do scary things," "explore strange places," "break rules") (Palmgreen *et al.*, 2001). Sensation seeking is included as a covariate to ensure that test and control groups are equated on it. Drug use could be directly measured so that the advertising effects could be assessed for drug user and nonuser subsamples, but this would add to the cost and sample size needs. Sample sizes are already very small for three-way interactions with just 50 participants per cell.

There is no correction for conducting multiple comparisons across the different belief and intent measures. The significance level is set at the standard $p < 0.05$, which means that on average 1 in 20 measures will show a significant effect due to chance alone, although this could be avoided by using a simple Bonferroni test (Winer *et al.*, 1991). Also, control participants do not see a control advertisement; they see nothing. It is generally recommended that control participants be given a pseudotreatment or placebo to avoid the Hawthorne effect in which being a full research participant itself improves responses. Specifically, a control advertisement such as a nondrug public service announcement could be used and control participants could answer questions about it (Andrews and Maronick, 1995, Pechmann and Reibling, 2006). However, the control advertisement must be truly neutral. Another issue is that some involved in the campaign think that an antidrug advertisement should be aired if it works on the target group, even if it may have adverse effects on a nontarget group. However, this seems problematic because often media spills over from one group to another. Finally, some involved in the campaign think advertisements that test poorly can be fixed and aired without retesting, but retesting it helps to ensure that problems are truly resolved.

Method 2

The US Federal Trade Commission (FTC) has employed copy testing in advertising deception cases for over 35 years (Maronick, 1991). Although it is not required per se, extrinsic evidence such as copy testing is usually considered at the FTC and given substantial weight in cases involving potentially deceptive implied ad claims (Andrews and Maronick, 1995). Extrinsic evidence can come in the form of methodologically sound consumer research studies, such as copy tests (Kraft, Inc., 1991), marketing research, and/or expert testimony.

Generally accepted copy test principles exist at the FTC, as well as substantial case precedent for FTC copy tests (e.g., Kraft, Inc., 1991, Stouffer Foods Corp., 1994; Thompson Medical Company Inc., 1984), but how these principles are operationalized may vary somewhat in each specific case (Maronick, 1991). Such generally accepted principles for copy testing should include the following: (i) the proper universe is examined, (ii) a representative sample is drawn from the universe, (iii) the mode of questioning is correct, (iv) the survey is designed by recognized experts, (v) the data gathered are accurately reported, and (vi) the sampling design, questionnaire, interviewing, and statistical analyses are in accordance with general accepted standards for conducting survey research (Manual for Complex Litigation, 1982).

Usually, an FTC advertising deception case begins with a complaint involving interstate commerce in the public interest (Andrews, 2001). Then, discovery is undertaken to obtain ads, media schedules, marketing research, and marketing plans; and an in-house copy test might be conducted. At a later point, the FTC might conduct an outside copy test with external marketing consultants if it appears that the case might go to trial. Typically, the copy test questions are pretested, followed by a main study with about 100 participants in the test group who see the challenged advertisement, and an additional 100 in the control group who see a control advertisement instead. Participants in the control group can be randomly assigned to see (i) a purged advertisement that is identical to the challenged advertisement, but with the challenged claims removed, (ii) a different advertisement for the same brand that does not contain the challenged claims (Kraft, Inc., 1991), or (iii) no advertisement, that is, a no-exposure control (Andrews and Maronick, 1995). Data collection is usually subcontracted to an experienced firm that uses research facilities in at least three geographically dispersed shopping malls.

A standard FTC copy test begins with screener questions to ensure that the sample is representative of the universe. In the case of Stouffer Foods Corp. (1994), the sample was drawn from a universe of principal food shoppers between the ages of 25 and 54 who had purchased a frozen entrée in the last three months and were not on a medically supervised diet. Once participants are successfully screened, they are shown the challenged advertisement or a control advertisement; it is shown with two clutter advertisements. Then, participants answer the survey questions, which generally follow a "funnel approach," beginning with general open-ended questions and then moving to successively narrower questions and ending with specific closed-ended ones (Kraft, Inc., 1991). It is important that the questions be related to the advertising claims; they cannot just measure general beliefs about a product attribute.

Hypothetically, an FTC case might be about a potentially misleading, low-fat advertising claim. After the respondents have finished looking at the target advertisement with clutter ads, a first set of questions would ask whether they remember seeing the advertisement and the brand in question. If answered correctly, they would be asked, "What does the brand's ad say or suggest to you about its products?" This would be followed with probes of "anything else" and would also include a "don't know" option. Next, respondents would be asked to read the target advertisement one more time. When finished, they would be queried, "Does the ad say or suggest anything about the amount of fat in the brand?" with the response options of yes, no, and, don't know/maybe/not sure. If the answer is "yes," they would be asked, "What does the ad say or suggest to you about the amount of fat in the brand's products?" with probes and don't know options. Respondents would then be asked, "Does the ad say or suggest that the brand's products are . . ." with the options presented as "low-fat foods," "high-fat foods," or "neither low- nor high-fat foods," as well as a "don't know" option.

Toward the end, other specific closed-ended and/or control questions may be presented. For instance, a control question is often asked involving an attribute that is not related to the challenged advertisement or the control advertisement either (e.g., about sugar in the case of challenged low-sodium claims, see Stouffer Foods Inc., 1994). The purpose of this is to help account for "yea-saying" on behalf of respondents. Lastly, demographic questions are collected. In the data analyses, the response percentages are computed, and then a net response is calculated by subtracting the responses to the control advertisement from the responses to the test advertisement. Sometimes, the respondent verbatims from open ends are summarized and presented along with the statistical results. Copy testing for corrective advertising is more involved. In the *FTC v. Novartis* case on Doan's Pills (Mazis, 2001), the tests assessed deception, claim materiality, false beliefs, ad-belief linkages, lingering beliefs, and remedies.

Strengths

The FTC's copy test method seems to comply with all 10 principles of good copy testing.

1. *Be relevant to the objectives*: The copy testing assesses the consumer's reasonable

interpretation and net impression of potentially misleading ad claims that are material or important (i.e., likely to affect conduct).

2. *Have agreement about the use of results*: The copy testing decision rules are consistent with generally acceptable principles, case precedent, and the 1983 FTC Deception Policy.
3. *Use multiple measurements*: A funneling process is used, from general open-ended questions to increasingly narrower questions and ending with specific closed-ended questions.
4. *Be based on a model of human response*: Copy testing is based on an information-processing model that recognizes the steps of exposure, awareness, comprehension, persuasion, and intent, though the focus is on comprehension (i.e., ad beliefs).
5. *Consider multiple exposures*: Participants see the test or control advertisement twice, once before general open-ended questions and a second time before more specific questions.
6. *Test comparably finished executions*: Advertisements are copy tested in final form, as they appeared in the media.
7. *Control the exposure context*: The test or control advertisement is shown with two clutter advertisements during the first exposure and alone during the second exposure.
8. *Define the relevant sample*: Screener questions are used to ensure that the sample is in the target market and demographics are assessed as well.
9. *Demonstrate reliability and validity*: Although single-item measures are commonly used, the funneling process allows for validity checks across measures. Also, control questions and don't know options help minimize the yea-saying bias.
10. *Take baseline (i.e., pre-exposure) measurements and/or use control groups*: Some participants are randomly assigned to see a control advertisement that does not contain the challenged claims.

Limitations

The FTC copy testing also involves tradeoffs. It can be difficult to examine low salience or tertiary claims if awareness of these claims is low in both the open- and closed-ended questions (Maronick, 1991). Also, on occasion, there could be strong halo effects; anything in the advertisement might be considered part of the challenged claims or might immediately elicit memories of a longer running campaign. This situation makes it difficult to create a control advertisement that excises or excludes the challenged claims (Andrews and Maronick, 1995). Researchers could use a no-ad-exposure control group, but the standard FTC questions are ad specific. Also, the FTC typically uses single item versus multi-item measures, which precludes standard reliability checks (i.e., alpha). However, the funneling procedure allows for some validity assessment across measures. Finally, most data is presented in the form of simple percentage comparisons between test and control groups or between test and control questions. Thus, there could be room for more sophisticated interval-scaled measures and analyses.

Method 3

The ARSgroup of The Quality Measurement Company has copy tested television advertisements for over 40 years. ARSgroup has conducted over 40 000 copy tests worldwide, often for major firms such as Campbell Soup, Procter and Gamble, and Wal-Mart. It specializes in gauging the effectiveness of television advertisements based on its ARS persuasion score. Its competitors include Nielsen IAG, The Pretesting Company, and Marketing Evaluations, among others. The information below is based on published articles about ARSgroup's advertising copy testing method (Adams and Blair, 1992; Blair and Rabuck, 1998; Stewart *et al.*, 1983). Overall, the ARSgroup's copy test research shows that advertising quality, as measured by their ARS persuasion score, is related to short-term (i.e., weekly or monthly) changes in brand sales and share. Also, advertising quality is more important than advertising expenditures in affecting sales and share. Finally, a high-quality advertisement builds sales quickly, but has diminishing returns over time because its effect gradually wears out.

Typically, the ARSgroup copy tests a television advertisement in four geographically

dispersed cities (Adams and Blair, 1992; Stewart *et al.*, 1983). Approximately 400–1000 randomly selected men and women are recruited to participate ostensibly to evaluate a television pilot and are instructed to report to a research facility on a specific date and time. As an incentive, participants are told that raffles will be conducted and that many of them will win free bags of products. When they arrive, participants are shown lists of brands and are asked to choose the brand they would most like to win in each product category. Next, a raffle is conducted and some people win the products they have chosen. Then, participants view a television program in which each test advertisement is embedded along with unrelated or noncompeting clutter advertisements. At the end, participants are shown the same lists of brands again, and are asked to choose the brands that they would most like to win this time. Again, a raffle is conducted and some people win the products they have chosen. For this reason, virtually all ARSgroup copy tests involve low-cost consumer packaged goods.

The raw ARS persuasion score is calculated as the percentage of participants who choose the test brand after exposure to the test advertisement, minus the percent who chose the test brand before exposure to the test advertisement. This raw score is then converted to an adjusted persuasion score; the ARSgroup has found that the brand's market share biases the copy test responses and so this bias is removed. Three days after the copy test, participants are interviewed by phone to assess advertising recall (Stewart *et al.*, 1983), but the main measure used to evaluate advertisements is the adjusted ARS persuasion score. The ARSgroup's reliability studies indicate the adjusted persuasion score's test–retest reliability is 0.93 (Stewart *et al.*, 1983). Its validity studies show the adjusted persuasion score predicts trial rates for new, frequently purchased, branded consumer goods ($r = 0.85$). ARSgroup has also validated its copy test method for established brands in split cable ad tests, meaning some households receive the advertisement while others do not, and household purchase rates are compared (Adams and Blair, 1992).

Strengths

The ARSgroup method generally complies with the 10 principles of good copy testing.

1. *Be relevant to the objectives*: The ARSgroup objective is to predict an advertisement's effect on brand sales and share; its validation studies indicate that it does this.
2. *Have agreement about the use of results*: ARSgroup's recommendations are clear – ads with low ARS persuasion scores should not be used because they will not improve sales no matter how much they air.
3. *Use multiple measurements*: The ARSgroup measures product choice as well as advertising recall.
4. *Be based on a model of human response*: There is no explicit model of advertising response.
5. *Consider multiple exposures*: Participants see an ad just once, but the ARSgroup's validation studies suggest this is sufficient (Adams and Blair, 1992; Blair and Rabuck, 1998).
6. *Test comparably finished executions*: The ARSgroup copy tests television ads in final or near final form; it does not test storyboards.
7. *Controls the exposure context*: The exposure context is carefully controlled. The advertisements are embedded in the same television program with the same clutter advertisements across multiple copy tests.
8. *Define the relevant sample*: A general audience sample is used, but most of the copy tests are for general or mass-market products.
9. *Demonstrate reliability and validity*: The ARSgroup has demonstrated reliability and validity.
10. *Take baseline (pre-exposure) measurements and/or use control groups*: The ARSgroup takes a pre-exposure measure of brand choice and uses this as the baseline for assessing the advertisement's persuasive impact on brand choice.

Limitations

The ARSgroup does not disclose the adjustments they make to the raw persuasion score to correct for response biases because of brand

market share; their adjustments are proprietary. This makes it difficult to fully evaluate the method and to publish ARSgroup studies; most academic journals require that measurement adjustments be disclosed. This proprietary approach also prohibits independent replication of the ARSgroup validation studies. In addition, many copy test participants receive the products they choose and so the advertised products must be low-cost consumer packaged goods. The ARSgroup method cannot be used for expensive items such as cars, or for illegal items such as marijuana. Since a representative sample is employed, the products must have reasonably wide appeal.

Areas of debate. There are many general areas of debate about copy testing. The main ones are (i) the reliance on convenience samples such as people in shopping malls, (ii) the artificial viewing environment including forced exposure to the advertising, (iii) no assessment of cumulative advertising effects, (iv) often no measure of actual behavior, (v) no assessment of advertising-induced word of mouth, and (vi) the increased use of on-line or web-based copy testing, which means less control over samples and exposure environments, although the effective use of screeners can help. A longstanding concern is the reliance on malls and other public venues that offer convenience samples. However, convenience sampling is common in marketing research and the US courts have accepted it (Maronick, 1991; Plevan and Siroky, 1991). The artificial viewing environment is of concern and so many copy tests include clutter. The FTC shows the test advertisement with two clutter advertisements and then measures awareness of it. The ARSgroup embeds the test advertisements with numerous clutter advertisements.

Copy tests typically offer no assessment of cumulative advertising effects, but most ensure that consumers have seen the test advertisement at least twice. Research suggests that an advertisement generally has its impact after just two or three exposures (Pechmann and Stewart, 1988). The ad's effect on actual product use or purchase behavior should be measured whenever possible, if this is relevant to the copy test objectives. The FTC copy tests often aim to measure false advertising beliefs. For other copy tests, behavior can often be measured if the budgets are adequate and the test advertisements promote low-cost consumer goods. For expensive items such as cars, illegal items such as marijuana, and high-risk items such as cigarettes, it can be very difficult to assess behavior. In such cases, it is important to use validated intent measures.

Copy tests generally do not measure word of mouth effects, for instance, when one person talks to another about a memorable advertisement. Word of mouth may produce either positive or negative outcomes. The ARSgroup's copy tests involve large groups of people who could be permitted to talk about the ads, but this would likely be too contrived. Internet postings may be a better way to assess word-of-mouth advertising effects. Also, a recent concern is that on-line or web-based copy testing might be less valid than traditional mall-based methods. People who participate in on-line copy tests may be less representative than people in malls, although effective screeners can help in this regard. Yet, the on-line environment can be less controllable than a mall facility; for instance, people on-line can talk to others and/or multitask while completing the survey. More research is needed to compare on-line versus mall-based copy tests.

In summary, copy testing is a standard and well-accepted marketing research method that has been shown to be reliable and valid if it is done well. Companies and other entities should feel confident about using copy testing to evaluate advertisements as long as the recommended procedures discussed above are followed.

Bibliography

Adams, A.J. and Blair, M.H. (1992) Persuasive advertising and sales accountability: past experience and forward validation. *Journal of Advertising Research*, **32** (2), 20–25.

Andrews, J.C. (2001) The use of marketing knowledge in formulating and enforcing consumer protection policy, in *Handbook of Marketing and Society* (eds P.N. Bloom and G.T. Gundlach), Sage Publications, Thousand Oaks, CA, pp. 1–33.

Andrews, J.C. and Maronick, T.J. (1995) Advertising research issues from FTC versus Stouffer Foods Corporation. *Journal of Public Policy and Marketing*, **14**, 301–327.

Belch, G.E. and Belch, M.A. (2007) *Advertising and Promotion: An Integrated Marketing Communications*

Perspective, 7th edn., McGraw-Hill Irwin, Boston, MA.

Blair, M.H. and Rabuck, M.J. (1998) Advertising wearin and wearout: ten years later–more empirical evidence and successful practice. *Journal of Advertising Research*, **38** (5), 7–18.

Clark Boardman Company Ltd (1982) *Manual for Complex Litigation*, Federal Practices Series, Clark Boardman Company Ltd., New York , NY.

Cook, T.D. and Campbell, D.T. (1979) *Quasi-Experimentation: Design & Analysis Issues for Field Settings*, Houghton Mifflin, Boston, MA.

Foley, D. and Pechmann, C. (2004) The national youth anti-drug media campaign copy test system. *Social Marketing Quarterly*, **10**, 34–42.

Kraft, Inc. (1991) Commission decision, 114 FTC 40-46, pp. 116–151.

Maholtra, N.K. (2009) *Basic Marketing Research: A Decision-Making Approach*, Prentice Hall, Upper Saddle River, NJ.

Maronick, T.J. (1991) Copy tests in FTC deception cases: guidelines for researchers. *Journal of Advertising Research*, **31**, 9–17.

Mazis, M.B. (2001) FTC v. Novartis: the return of corrective advertising? *Journal of Public Policy & Marketing*, **20**, 114–122.

PACT (1982) Positioning advertising copy testing. *Journal of Advertising*, **11** (4), 4–29.

Palmgreen, P., Donohew, L., Lorch, E. *et al.*, (2001) Television campaigns and adolescent marijuana use: tests of sensation seeking targeting. *American Journal of Public Health*, **91** (2), 292–296.

Pechmann, C. and Reibling, E. (2006) Antismoking advertisements for youth: an independent evaluation of health, counter-industry, and industry approaches. *American Journal of Public Health*, **96** (5), 906–913.

Pechmann, C. and Stewart, D.W. (1988) Advertising repetition: a critical review of wearin and wearout, in *Current Issues and Research in Advertising* (eds J.H. Leigh and C.R. Martin), University of Michigan Press, Ann Arbor, MI, pp. 285–330.

Plevan, K.A. and Siroky, M.L. (1991) *Advertising Compliance Handbook*, 2nd edn., Practicing Law Institute, New York, NY.

Shimp, T.A. (2010) *Advertising, Promotion, and other Aspects of Integrated Marketing Communications*, 8th edn., South-Western Cengage Learning, Mason, OH.

Stewart, D.W., Furse, D.H., and Kozak, R.P. (1983) A guide to commercial copytesting services. *Current Issues and Research in Advertising*, **6** (1), 1–43.

Stouffer Foods Corp. (1994) Commission decision, 118 FTC pp. 746–820.

Thompson Medical Company, Inc. (1984) Commission decision, 104 FTC pp. 648–844.

U.S. Federal Trade Commission (1983) Policy Statement on Deception October 14, 1–19; appended to *Cliffdale Associates, Inc*. (1984), Commission Decision, 103 FTC 174–184.

Vakratsas, D. and Ambler, T. (1999) How advertising works: what do we really know? *Journal of Marketing*, **63**, 26–43.

Winer, B.J., Brown, D.R., and Michels, K.M. (1991) *Statistical Principles in Experimental Design*, 3rd edn., McGraw-Hill, New York, NY.

sex in advertising

Tom Reichert

Few topics in marketing prompt as much interest as sex in advertising. Salacious ads can generate hundreds of blog posts and coverage from traditional media such as CNN and ABC News. Sometimes the advertisement leads to increased sales, and other times it results in consumer complaints and a swift apology as the ad is promptly pulled off the air. Regardless of the result, advertisers have used sexual imagery and themes to promote their products since the earliest days of marketing and will continue to do so in a quest for sales.

Definition

Sex in advertising is often defined as "sexuality in the form of nudity, sexual imagery, innuendo, and double entendre . . . employed as an advertising tool for a wide variety of products" (Courtney and Whipple, 1983, p. 103). Research indicates that the vast majority of sexual content in advertising consists of images of sexually appealing models – models who adhere to ideals of Western beauty standards (Reichert, 2003). Often the models are clothed in a revealing manner and posed or behaving – either by themselves or with other models – to stimulate sexual thoughts and feelings within viewers (see Figure 1). Lighting, pacing, editing, setting, and other aspects of the prevailing tone contribute to the sexual nature of an advertisement. Verbal elements also contribute to sexual meaning such as sexually suggestive phrases and intonation, or headlines with double meaning that may seem innocuous until paired with a provocative image.

Sexual content within an ad may or may not be relevant to the product category or advertised brand message. For example, some ads feature images of scantily clad women or men making eye contact with the viewer but with no logical connection to the advertised product. A Carl's Jr. commercial aired in 2005, for instance, featured a swimsuit clad Paris Hilton seductively washing a Bentley and taking a bite from a hamburger. More often, however, sexual content is central to the brand's message. For example, advertising can tell consumers that the brand will enhance

Figure 1 Nudity is clearly present in this marketing image for youth clothier Abercrombie & Fitch. (reproduced by permission of Sam Shahid).

their beauty or chances for romantic success. Commercials aired in the 1970s for Prince Matchabelli's Wind Song perfume positioned the fragrance as a romantic attractant with the memorable jingle, "I can't seem to forget you. Your Wind Song stays on my mind." In a typical commercial, a women sprays a napkin with Wind Song for the waiter to deliver to her boyfriend across the restaurant. Immediately, her boyfriend appears and bestows a romantic kiss. Most recently, men's toiletry brands such as Axe and Old Spice position their body sprays and deodorants as a key to attracting females.

In the late 1960s, there was growing concern regarding subliminal uses of sex in advertising. Advertisers were accused of inserting virtually undetectable phallic symbols or words such as "SEX" into ads that are processed by the sub- or preconscious mind. Following exposure, consumers find themselves drawn to these brands without realizing the connection to hidden sexual content. Much of the interest in this topic is attributable to Key's (1973) best-selling book, *Subliminal Seduction*. Advertisers counter that they can hardly gain consumer attention by using readily perceptible words and images, much less hidden symbols. Marketers also argue that at no point in the ad-creation process is a request made to insert subliminal sexual content.

Effectiveness

Since the 1960s, academic researchers have sought to determine if and how sexual content works when used in advertising. Overall, findings indicate that sex in advertising is

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

particularly adroit at getting noticed (Belch *et al.*, 1981; Percy and Rossiter, 1992). In other words, sex in advertising achieves a high level of attention. A distraction effect can occur, however, such that attention is directed toward the sexual content instead of the brand message or sponsorship information (Severn, Belch, and Belch, 1990). As a result, sexual ads get noticed and are likely to be remembered, but consumers may not recall what was said or what brand was advertised.

Findings also reveal that sex in advertising generates an emotional response within viewers (LaTour and Henthorne, 1993). Arousal is more pronounced as the explicitness of the sexual content (i.e., nudity) increases. Similarly, explicitness is also linked to the degree a viewer likes or dislikes an ad. Recent research shows that some personality characteristics such as sensation seeking and sexual schemas influence how a consumer will respond to sex in advertising (Reichert, LaTour, and Kim, 2007).

While some advertisers only use sex to grab attention, many use it as a selling point for their brand. One study found that most examples of sex in advertising contain a brand-related message that buying the brand will result in the purchaser being more sexually attractive, getting more sex or better sex, or experiencing sexy feelings that will boost one's self confidence (Lambiase and Reichert, 2003). Any distraction effect may be mitigated if one of these sexual brand promises is a central theme of the ad. However, credibility – whether consumers actually believe the advertiser's promises – is important to consider as well.

As described previously, relevance between sex and the advertised product influences the effectiveness of the ad. Research typically shows an advantage in terms of memory, likability, and purchase intention when sex in advertising is used for products with a reasonable connection to sex (Percy and Rossiter, 1992). Typical product categories generally viewed as relevant to sex in advertising include fashion, fragrance, tobacco, and alcohol. A logical connection can be made to other products or brands but effectiveness is reduced, the further the connection between the product and sex.

While somewhat stereotypical, women and men often respond differently to sex in advertising (Reichert, LaTour, and Kim, 2007). For example, whereas men are less likely to be turned off or offended by a gratuitous ad, there is evidence that women respond more positively to images of couples and romantic settings. Perhaps, not surprisingly, both women and men respond more positively to images of the opposite sex compared to images of the same sex.

History and Prevalence

Early uses of sex in advertising date back thousands of years (Reichert, 2003). Painted signs for brothels intended for weary sailors have been discovered in port towns along the Mediterranean. As early as the 1800s, classical images of nude or semiclothed women appeared in posters and handbills designed to attract the attention of male patrons. In 1912, the J. Walter Thompson (JWT) agency claims to have invented the "sexual appeal" in ads for Woodbury's Facial Soap. Its ads featured beautiful women accompanied by adorning men. The tagline, "The skin you love to touch," made it clear that buying and using Woodbury's soap was the key to love and romance. Whether JWT truly invented the sexual appeal or not, such appeals – the promise of attracting a mate – continues to be a commonly employed appeal for many types of socially consumed products.

In the 1960s, sex in advertising migrated from the pages of men's and women's magazines to network television. For example, a commercial for Noxema shaving cream aired in 1966 featured Gunilla Knutson, a former Miss Sweden, telling men to "Take it off. Take it all off," while "The Stripper" score played in the background.

Calvin Klein fashion ads generated media attention from 1979 to 1995, and today, for sexual ad content that often crosses the line of social acceptability (Gaines and Churcher, 1994). For example, commercials in 1981 featured a then 15-year-old Brooke Shields teasingly uttering double entendres such as, "Know what comes between me and my Calvins? Nothing." Klein subsequently introduced his men's and women's intimate-wear lines with beautiful models clad only in underwear – not as

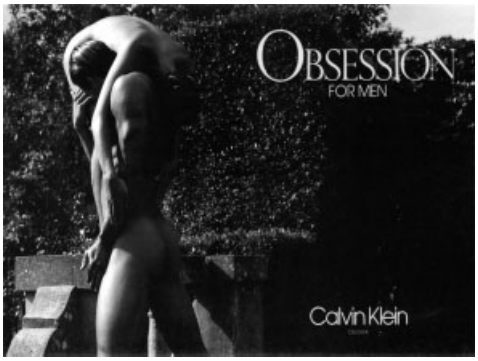


Figure 2 In the 1980s, nudity in Calvin Klein ads pushed the limits of acceptability in the United States. (reproduced by permission of Sam Shahid).

common an occurrence as in today's advertising – in huge outdoor ads in Times Square. His ads in the 1980s and 1990s for his designer fragrances, especially Obsession, featured nude models in idyllic scenes (see Figure 2) or in steamy ménage-a-trios settings. And in 1995, Klein again generated controversy with a campaign featuring underage models filmed in what appears to be a basement. Many critics claimed that the commercials contained a pedophilic theme. Although controversial, over the years, the ads and the attention Calvin Klein garnered often resulted in at least a short-term sales increase.

Today, sex in advertising is relatively common. Studies indicate that up to 10% of network and cable television commercials contain sexual content (Hetsroni, 2007). To lure viewers, up to 25% of network promos – the ads networks air to promote their programming – contain sexual scenes. Overall, sexual themes and images are more likely to appear in men's and women's fashion, beauty, and general-interest magazines (Soley and Reid, 1988). Online ads contain sexual content as well; up to 20% of banner ads on the most popular news and sports sites contain images of physically attractive models. In addition, websites can contain sexual images on the site or within on-line video. For example, in 2009, Hollister, a clothing brand for teens and young adults, featured partial nudity and intimate scenes in streaming video for its brand of intimate wear. It is doubtful that the video would have cleared any mainstream TV network's standards board.

Criticisms

Sex in advertising generates criticism within the profession and by social critics. Generally, advertising professionals do not respect the use of sex in advertising. Many creative people within the advertising business consider its use a lazy way to grab attention. David Ogilvy said it this way: "Some copywriters, assuming that the reader will find the product as boring as they do, try to inveigle [the viewer] into their ads with pictures of babies, beagles and bosoms. This is a mistake" (Ogilvy, 1985, p. 139). Industry professionals believe that advertising should appeal to people's higher motivations and attempt to uplift people, not speak to their lowest, most base, motives.

Socially, sex in advertising is criticized because of its slanted depiction of women. While some ads feature "beefcake" images of men or healthy images of couples, most sexual ads feature images of women portrayed as objects of desire (Hetsroni, 2007; Soley and Reid, 1988). Feminists object to these one-dimensional images because they contribute to the impression that a woman's value resides in her physical attractiveness and ability to attract a man. A related concern is that pornographic themes such as bondage and related fetishes are showing up in mainstream advertising. Pornographic or not, many people consider sexual content offensive and intrusive, and resent seeing it in a commercial when viewing television with the family – especially when children are present.

Despite these criticisms, some marketers continue to employ sex in advertising to get their products noticed and to convince consumers that their brands can enhance consumers' relational intimacy and satisfaction.

See also *advertising message appeals; luxury brands versus other brand categories; subliminal advertising; the role of signage in marketing: outdoor advertising, out-of-home media, and on-premise signs*

Bibliography

Courtney, A.E. and Whipple, T.W. (1983) *Sex, Stereotyping, and Advertising*, D.C. Heath, Lexington.

Belch, M.A., Holgerson, B.E., Belch, G.E., and Koppman, J. (1981) Psychophysical and cognitive responses to sex in advertising, in *Advances in Consumer Research*, Vol. 9 (ed. A. Mitchell), Association for Consumer Research, Ann Arbor, pp. 424–427.

Gaines, S. and Churcher, S. (1994) *Obsession: The Lives and Times of Calvin Klein*, Avon, New York.

Goodrum, C. and Dalrymple, H. (1990) *Advertising in America: The First 200 Years*, Abrams, New York.

Hetsroni, A. (2007) Sexual content on mainstream TV advertising: a cross-cultural comparison. *Sex Roles*, **57**, 201–210.

Key, W.B. (1973) *Subliminal Seduction: Ad Media's Manipulation of a Not So Innocent America*, Prentice-Hall, Englewood Cliffs.

Lambiase, J. and Reichert, T. (2003) Promises, promises: exploring erotic rhetoric in sexually oriented advertising, in *Persuasive Imagery: A Consumer Perspective* (eds L. Scott and R. Batra), Erlbaum, Mahwah, pp. 247–266.

LaTour, M.S. and Henthorne, T.L. (1993) Female nudity: attitudes toward the ad and the brand, and implications for advertising strategy. *Journal of Consumer Marketing*, **10**, 25–32.

Ogilvy, D. (1985) *Ogilvy on Advertising*, Vintage, New York.

Percy, L. and Rossiter, J.R. (1992) Advertising stimulus effects: a review. *Journal of Current Issues and Research in Advertising*, **14**, 75–90.

Reichert, T. (2003) *The Erotic History of Advertising*, Prometheus, Amherst.

Reichert, T., LaTour, M., and Kim, J.Y. (2007) Assessing the influence of gender and sexual self-schema on affective responses to sexual content in advertising. *Journal of Current Issues and Research in Advertising*, **29**, 57–71.

Severn, J., Belch, G.E., and Belch, M.A. (1990) The effects of sexual and non-sexual advertising appeals and information level on cognitive processing and communication effectiveness. *Journal of Advertising*, **19**, 14–22.

Soley, L. and Reid, L. (1988) Taking it off: are models in magazine ads wearing less?. *Journalism and Mass Communication Quarterly*, **65**, 960–966.

viral marketing on the Internet

Petya Eckler and Shelly Rodgers

Definition

Viral marketing applies traditional word-of-mouth (WOM) marketing to the online environment. Much confusion exists about its actual definition, as the terms viral marketing, stealth marketing, buzz marketing, and viral advertising have been used interchangeably (Golan and Zaidner, 2008). The authors define *viral marketing* as a broad array of online WOM strategies designed to encourage both online and peer-to-peer communication about a brand, product or service (p. 961). These include, but are not limited to, viral videos, email messages, use of online social networks and online forums, text messaging, interactive microsites and online games, blogs, podcasts, and so on. *Viral advertising*, on the other hand, is a subset of this larger umbrella and is defined as unpaid peer-to-peer communication of provocative content originating from an identified sponsor using the Internet to persuade or influence an audience to pass along the content to others (Porter and Golan, 2006).

History

The term *viral marketing* is credited to venture capitalists Steve Jurvetson and Tim Draper who used viral to describe the marketing techniques behind Hotmail's growth in 1996. The catalyst for success was the promotional tag on each outgoing email, which turned subscribers into company salespeople who took the message to their own WOM networks (Jurvetson and Draper, 1997). Thus, the viral message spread organically with spatial and network locality, much like a virus, and garnered 12 million subscribers in 1.5 years for less than $500 000 (Jurvetson and Draper, 1997). Since then, viral campaigns have become increasingly popular.

Unique Features

Unique to viral marketing is its focus on the message, not the product. Thus, the brand is often secondary in viral videos, as they aim to look more like entertainment pieces and less like branded commercials. The success of a viral campaign can be attributed to its emotional or entertainment value rather than information about the brand or product (MindComet, 2006). Appeals related to humor, sexuality, nudity, and violence are more common as compared to television ads (Porter and Golan, 2006). More extreme appeals are often needed because viral ads must prompt "forwarding" behavior not just awareness or liking, as with traditional advertising.

In viral campaigns, initial exposure is achieved through seeding, which "plants" the message in selected consumers, who then spread it to their social networks. Seeding decisions depend on the target audience for each campaign (MindComet, 2006).

Similar to other marketing tools, viral marketing has unique benefits and drawbacks. Advantages include the following:

- reduced cost of promotion because of free peer-to-peer distribution and eliminated media buys;
- increased credibility due to friends' endorsement (Chiu *et al*., 2007);
- increased visibility as messages cut across media clutter (MindComet, 2006);
- decreased interruption as viewers choose the time and place to view viral messages (MindComet, 2006); and
- improved format flexibility as messages can take various forms.

Viral marketing disadvantages include the following:

- reduced control of the marketer, as a viral campaign relies on consumers to spread it;
- increased reliance on consumers' motivation to spread the message for campaign success;
- increased risk of negative reactions as viral ads may become too unconventional.

Effective viral marketing depends on consumers' forwarding behavior; therefore, understanding consumer motivations is vital, as is research on the content features that motivate consumers and the consumers most appropriate for initial seeding. As with other marketing tools, viral approaches must be integrated into the

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

MARKETING MIX and typically are not a standalone promotional strategy.

Bibliography

Chiu, H.-C., Hsieh, Y.-C., Kao, Y.-H., and Lee, M. (2007) The determinants of email receivers' disseminating behaviors on the Internet. *Journal of Advertising Research*, **47** (4), 524–534.

Golan, G.J. and Zaidner, L. (2008) Creative strategies in viral advertising: an application of Taylor's six-segment message strategy wheel. *Journal of Computer-Mediated Communication*, **13** (4), 959–972.

Jurvetson, S. and Draper, T. (1997) Viral Marketing Phenomenon Explained, http://www.dfj.com/news/article_26.shtml (retrieved 13 February 2009).

MindComet (2006) Viral Marketing: Understanding the Concepts and Benefits of Viral Marketing, White paper.

Porter, L. and Golan, G. (2006) From subservient chickens to brawny men: a comparison of viral advertising to television advertising. *Journal of Interactive Advertising*, **6** (2), 26–33. http://www.jiad.org/article78 (retrieved 13 February 2009).

advertising media selection and planning

Marjolein Moorman, Peter C. Neijens, and Edith G. Smit

Introduction

Media are essential in the communication of brands with their target groups. After all, a splendid commercial is ineffective if it does not reach its audience. A media strategy for a brand involves the selection of media types (e.g., newspapers, television), vehicles (e.g., Home and Garden, Friends), and media units (e.g., a full-page, full-color, back page ad in the December issue). Other subjects that should be dealt with in the media plan are timing (when a message should be communicated) and frequency (how often).

Media (any transmission vehicle or device through which communication may occur) can take many forms. In a broad sense media include not only traditional (mass) media, such as broadcast media (TV and radio), print media (newspapers and magazines), out-of-home media (posters, billboards), new media (Internet, games, mobile phone), but also events, sponsored persons, retail outlets, and sales persons. This contribution focuses on traditional and new media.

Media decisions are important as media are by far the biggest debit entry in a campaign. According to a study by the Advertising Research Foundation (2002, cited in Kelly and Juggenheimer, 2004, p. 4), not less than 80–85% of the costs of an average campaign is spent on buying advertising space. The other 15–20% is used for research, development of the commercial, production, and evaluation.

Media decisions have become more complicated, more extensive, and thus more interesting in the last decades because of considerable changes in the media landscape. These changes include an enormous growth in the media offer, in traditional media as well as new media such as the Internet, mobile phones, and games. This growth gives unprecedented opportunities, but poses problems as well because target groups are fragmented across media. For example, in 1995 there were 225 shows in British television that reached audiences of over 15 million, but by 2005 there were none (FEN (Future Exploration Network), 2007, p. 4). Another problematic aspect of the new media landscape is the advertising clutter that has grown to unprecedented levels. For instance, today an average television viewer in the Netherlands is exposed to 17 minutes of daily commercial viewing, compared to 5 minutes in 1990. The advertising clutter makes it more difficult for advertisers to attract attention to their messages and generate advertising effects. Another factor that complicates the life of media planners is the shift in control over brand communication from advertisers to consumers, instigated by new, interactive media.

In this article we discuss the factors that play a role in the selection and planning of media in a campaign. First, we discuss key concepts (reach, frequency, timing, and costs). Then we pay attention to audience and medium factors that influence media impact. Finally, we discuss developments in new and traditional advertising media (*see* BUSINESS-TO-BUSINESS MEDIA SELECTION).

Key Concepts

Reach. The basis currency in media planning is the number of people being *reached* by an advertising message carried by a particular media vehicle. For broadcast media, the term *ratings* is used. It is common to make a distinction between different indicators of reach (see Harvey, 1997) such as vehicle distribution, vehicle exposure, advertising exposure, ad attentiveness, ad communication, ad persuasion, ad response, and sales response. Advertising exposure – "the number of target audience individuals exposed to the advertising" (Rossiter and Percy, 1998, p. 447), also defined as *the number of people with open eyes/ears in front of advertising space* – is generally considered to be the most valid indicator for reach. Other frequently used indicators such as vehicle distribution and vehicle exposure are less accurate estimations of advertising reach. Moreover, "deeper concepts" such as persuasion and behavioral responses are not appropriate because these are affected by factors beyond the control of the media, such as the power of the ad and the attractiveness of the bargain offered in the ad (*see* ADVERTISING EFFECTIVENESS).

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Media planning terminology contains a variety of reach-related concepts. *Current reach*, for example, indicates the number of people reached within the publication interval of a title (a day for dailies, a week for weeklies, etc.). Another concept is *selectivity*, which expresses the number of target group people that are reached compared to the total number of people reached. The higher the selectivity of a medium, the lower the *waste* (i.e., reach of people that do not belong to the target group). A third example in this respect is *effective reach*. This concept indicates the number of target audience individuals reached at the effective frequency level (is dependent on a particular campaign, e.g., three exposures or more; see section Frequency below). Other metrics include, for example, *ad impressions* for interactive advertisements (see www.iab.net).

Advertising reach is rarely evaluated on the level of an individual message, but is done mostly on the level of the campaign as a whole. Campaign reach indicates the number of *contacts* with all the published advertising messages in the campaign. To illustrate this concept: if one person is exposed to three commercials in a row, the campaign has generated three contacts. We have the same number of contacts if three people are exposed to one commercial each. We call the total number of contacts with a campaign *gross reach*. The *net reach* in a medium plan is the number of people that are exposed to the campaign at least once. In other words, gross reach minus the *overlap* (i.e., repeated exposures). In both examples, the gross reach of the campaigns is 3, but the net reach is 1 and 3, respectively.

For broadcast media, gross reach is usually expressed in *Gross Rating Points* (GRPs). A value of 100 GRPs means that the campaign has a number of contacts that is equal to 100% of the target audience. That does not necessarily mean that all target group members are reached. It may mean that 50% of the target group is reached twice, in other words: 50% of the target group has a (average) contact frequency of 2. GRPs are an estimate of the total number of exposures per 100 target audience members in an advertising cycle, without taking into account if these are unique or repeated exposures. We can express this in a simple formula: GRPs = net reach × average frequency. This formula illustrates that a media planner has to make a trade-off between "reach" and "frequency," because the same budget can be used to reach as many people as possible ("how many"), or used to confront a (smaller) audience as often as possible ("how often").

It may be noted that a campaign generates a frequency distribution of contacts: in a campaign that aims to reach people three times on average, some individuals will have one contact with the campaign, others two, five, or seven, and so on. That complicates the decisions of the media planner. The size of the overlap between the reach of the various medium vehicles in a campaign determines to what extent more insertions of a particular medium vehicle lead to a higher reach or to a higher contact frequency.

Frequency. How often should people be reached with a message? Decisions with respect to frequency of exposures are important because media costs are high and campaigners want to avoid wasting money due to too little insertions (lack of effects) or too many insertions (unnecessary costs). Most theories that specify a relationship between frequency of exposures and campaign effects (the so-called response curve) assume that repeating the message is useful because it adds to the effects of the campaign (wear-in). After a certain number of repetitions the effects decline, or can even become negative (wear-out). However, theories differ strongly with respect to the supposed shape of the response curve and the optimal frequency. The frequency debate was especially vigorous in the last three decades of the twentieth century. We will discuss a few well-known frequency theories from that time in chronological order.

One of the most influential theories on the optimum number of exposures is the *three-hit theory* of Herbert Krugman. According to Krugman (1972), a cognitive reaction dominates the response of the audience during the first confrontation with a commercial. The receiver asks himself: "What is it? What is it all about?" During the second confrontation, an evaluative reaction dominates: "What is in it for me?" The third confrontation is the real reminder. The viewer knows what

it is all about and can take action. Krugman discusses *proven exposures*. According to him, it is possible that a commercial brings the receiver to his second response ("What is in for me") after, for example, the twenty-third confrontation.

Naples (1979) presented a literature review on the authority of the American *Association of National Advertisers* (*ANA*) entitled *Effective Frequency: The Relationship Between Frequency and Advertising Effectiveness*. His conclusions were: "One confrontation with an advertisement has no or little effect, except in extraordinary circumstances. Two confrontations (within a purchase interval) is effective. The *optimal* frequency is at least three confrontations within a purchase interval".

On the basis of analyses of Nielsen-panel television viewing and buying behavior data, John Philip Jones concluded in 1995 that the first blow is half the battle. He stated that a successful campaign is effective from the first confrontation. In other words, the response curve has no threshold. He also concludes that the first confrontation will have the greatest effect and that the next confrontations will have less effect (diminishing returns): the response curve is convex. Critics of Jones have brought forward that he had only looked at short-term effects of established brands in the fast-moving consumer goods (FMCG) sector and that his conclusions therefore had limited validity.

McDonald (1996) wrote a follow-up of Naples' review and came to other conclusions. He subscribed to Jones' ideas with respect to the short-term effects of advertising for established brands. However, he concluded that new messages, for new brands, do show a threshold. Furthermore, the shape of the response curve depends, according to McDonald, on a number of factors. In other words, there is not one rule only for optimal frequency. The following factors have an influence on the shape of the response curve: activities of competitors, share of voice, market share and status of the brand, the quality of the advertisement, the time until the purchase, the target group, and the type of product.

Timing. Timing is another important media planning decision, which relates to the length and moment of the campaign period. A choice has to be made between "bursting," flighting, or pulsing (large media exposure over a short period) on the one hand and "dripping" (spreading small exposure over time) on the other hand (Heflin and Haygood, 1985). Heflin and Haygood concluded that media schemes with an average concentration showed the best results. They explain the bad results of the high intensive schemes because these irritate and satiate the audience. The bad results of the very low intensive schemes were explained by cognitive constraints ("forget effect"). According to Ligthart (1999), the answer to the question "*bursting or dripping?*" depends on a number of factors. First, if a large audience can be reached quickly, then a burst is preferred. Second, dripping is more effective for messages that are forgotten quickly because, for example, many competitors are advertising, the ad stock is low, or the quality of the commercial is poor.

A final factor is the budget. When the budget is substantial, an effective combination of bursting and dripping ("maintenance") is possible (Ligthart, 1999, p. 56).

Other timing decisions concern the period of the year in which to advertise: the season, the week, the day, and the time of the day. Moreover, advertisers have to decide when to advertise in relation to competitors in the market: simultaneous or not. An important criterion here is the possibility to get a high *share of voice*, which is important for the visibility of the brand. A high share of voice depends on the relative amount of money that an advertiser can spend.

Cost. In the selection of media for a campaign, costs play an important role. Several metrics are used to express the costs: absolute cost, cost per 1000 persons reached (CPM), cost per GRP (CPP, cost per rating point). In the Internet context, variants include cost per click through (CPT or CPC), cost per lead (CPL), cost per sale (CPS), cost per impression (CPI), pay per sale (PPS), or pay per purchase (PPP). Rate cards usually mention gross costs. Big advertisers and media agencies are often able to negotiate substantive discounts (net cost) (*see* COMMUNICATIONS BUDGETING; THE ADVERTISING BUDGET).

Impact Factors

Media differ in their ability to get a message across. The impact of the message may differ as a consequence of medium characteristics or audience reactions toward the medium.

Medium characteristics. First, media differ with respect to the number and type of modalities such as text, audio, pictures, and video. Television, for example, uses video, colors, and sound. This makes the medium highly suitable for the communication of emotions and the demonstration of products. Obviously, a billboard is much more restricted in this respect. Other characteristics that influence the impact of a medium are, for example, the format and the use of color.

Second, we can make a distinction between push media (also called display or delivery media) and pull media (also called search or retrieval media) (Dijkstra, Buijtels, and Van Raaij, 2005). Push media, such as television or radio, are media where the sender determines what is on offer and when. Pull media, such as print media or the Internet, are media where the receiver actively selects the message. These media require active and involved participants. Push media may be better suited for products and brands that are not (yet) in the consideration set of the consumer, or at the top of the response hierarchy.

Third, media also vary to the extent to which the audience can influence the pace of the information flow. *Internal pacing* (the receiver determines) is high for magazines and newspapers, and low for radio and television (external pacing, the medium determines). A medium with internal pacing is better suited for complex messages and for products and services for which consumers actively search information, because the consumer can adjust the pace of the information consumption to the complexity of the message.

Fourth, the amount of reach and the speed of reach accumulation differ between types of media. Television and out-of-home, for example, are characterized by the fast and large accumulation of reach, in contrast to, for example, magazines.

Audience responses. Audiences react differently to different media. Krugman (1965), for example, describes print media as "hot" media, requiring active engagement of the readers, while he considered broadcast media as "passive" or "cool" media with less involved audiences.

Audience involvement may not only differ between medium types, but also between media vehicles. This aspect is, amongst others, studied in so-called medium context studies. The medium context concerns the environment of the ad provided by the vehicle carrying it, such as a television program, an issue of a magazine, or an Internet site. A distinction can be made here between editorial context and commercial context. Studies on commercial context have predominantly concentrated on the effect of the amount and nature of other commercial messages in the environment of an ad, referred to as *clutter* and *competitive clutter*. It has been shown that as the number of ads in the environment of the target ad increases the effectiveness of the target ad decreases, especially when the other ads are directly competitive (Kent, 1995).

Studies on editorial context have shown time and again that the same source delivering the same message to the same audience on separate occasions produces different effects depending on the differing programming or editorial contexts in which the message appears (Norris and Colman, 1992). In particular, context-induced psychological responses, such as involvement elicited by a documentary, happiness caused by a sitcom, or sadness generated by a drama series, are considered to have an important impact on advertising processing (Moorman, Neijens, and Smit, 2005). At the heart of the theoretical explanations for the influence of context-induced psychological responses on advertising processing lies the assumption that mental reactions toward the editorial context do not cease when the editorial content is interrupted by advertisements, but that these reactions "carry over" to the advertisements. These carried-over reactions, in turn, influence advertising processing. For example, it has been frequently found that advertisements placed in editorial contexts which induce positive feelings are evaluated more positively. Furthermore, it has been shown that context-induced involvement has an influence on attention for advertisements and memory for advertisements.

Attitudes of media users toward the medium (vehicle) and advertising in the medium are also a source of influence. Media vehicles may be seen by their viewers or readers as authoritative, credible, reliable, or undependable. Advertising in some media vehicles, for example, in do it yourself programs or magazines may be appreciated more than in others. Some media vehicles attract audiences that are ad lovers, and others have a greater group of ad avoiders (Smit and Neijens, 2000) (*see* ATTITUDES).

Situational circumstances in which a person is exposed to an advertisement are also important for the impact of a medium. This includes the person's physical environment (e.g., "at home, at the kitchen table"), the social environment (e.g., "in the company of three family members"), the time frame ("during breakfast"), and the mental state a person is in prior to exposure to the medium content (e.g., "an early morning mood"). Although it has been shown that each of these aspects can influence advertising effectiveness substantively, receiver context aspects are less relevant for media planning because planners have no ways of influencing them.

New Developments

New developments in advertising media can be characterized as more multimedia, targeted, customized, and personalized as well as integrating advertising and transaction.

Multimedia. Many campaigns make use of more than one medium. In these cross-media or cross-tools campaigns, marketers seek to maximize the effectiveness of their budgets by exploiting the unique strengths of each medium and tool and by maximizing cross-media consistency and synergies. As Ephron (2000) stated: "Old media planning was about picking individual media. New media planning is about picking combinations of media (and permutations of media, where sequence of exposure is important)." (*see* INTEGRATED MARKETING COMMUNICATION STRATEGY; MULTICHANNEL MARKETING; INTEGRATED MARKETING COMMUNICATION).

One of the biggest challenges in advertising research these days is finding an answer to the question as to how each medium in a multimedia campaign contributes to the campaigns impact. An answer to this question is not easily found. Data from standard industry audience measurement studies do not offer a sufficient answer, because these studies are traditionally focused on measuring just one type of medium. This inhibits insight into the overlap between people's use of various media. Consequently, we cannot learn from these data which medium has produced the biggest effect.

A related, perhaps even more challenging, question is how each medium in a multimedia campaign enhances the contribution of other media. This question is driven by "the potential existence of synergy, that is, the added value of one medium as a result of the presence of another medium, causing the combined effect of media to exceed the sum of their individual effects" (Naik and Raman, 2003, p. 385). The question of synergy is, in other words, whether the combined effect of multimedia activities, such as television, radio and print, is greater than the total sum of their individual effects.

Various theories have been put forward in literature (e.g., Dijkstra, Buijtels, and Van Raaij, 2005) with respect to which synergetic effects from the use of multiple media in an advertising campaign can be expected. First, the *encoding variability theory* suggests that when a message is received through a variety of media, the information from the message will be encoded in memory in a more complex fashion than if the message was only received through one medium. The resulting information network in the brain will consequently be stronger, clearer, and more accessible, which enhances the likelihood that the information will be recalled accurately. Second, the *credibility principle* suggests that a message is perceived as more believable when it is received through multiple independent sources. Third, the *transfer hypothesis* proposes that the perception of an advertisement in one medium may be enriched by a previous encounter with the advertisement in another medium. For example, the images seen in a television commercial may be brought to mind when one hears the brand's jingle in a radio commercial (visual transfer), or the brand's jingle may be remembered when one sees a billboard along the highway (audio transfer). Or the other way around: a second confrontation with a message enriches the first impression (backward transfer).

Fourth, based on the *complementary principle* it can be expected that "synergy occurs when different media with their specific strengths complement each other in a campaign or when the strength of one medium compensates for the weakness of another medium" (Dijkstra, Buijtels, and Van Raaji, 2005).

On the basis of these four theories, it can be expected that multimedia campaigns are more effective than single media campaigns. It is not surprising therefore that most practitioners and academics share the belief that using multiple media tools is efficient, rather than extravagant or redundant. However, empirical evidence supporting this contention has been scarce until date.

New media. New media include digital media such as the Internet, mobile phones, digital games, and so on. These media offer a large number of advertising possibilities varying from the more "traditional" display ads (such as online banner ads, bumper ads before movies, e-mails) to advertising forms that reflect the new possibilities of digital media. New media's interactivity makes it possible to tailor advertising to the individual consumer, and to integrate transactions with communication. Search engine ads, for example, are ads shown in connection with search engine results. These ads are often as relevant for the consumer as the content they are searching for (*see* DIRECT AND INTERACTIVE MARKETING; E-COMMERCE AND INTERNET MARKETING).

Digital media make it also possible for audiences to create and communicate brand information on a large scale. With the emergence of consumer-generated media platforms, word-of-mouth conversations have migrated to the World Wide Web, creating a wealth of product-related information, often articulated in the form of online consumer reviews. These online consumer reviews are assumed to be more influential than marketer-generated information in determining consumers' product and brand attitudes (Park, Lee, and Han, 2007). The persuasive impact of consumer reviews, as well as other forms of (electronic) word of mouth, is often attributed to the perceived noncommercial nature of the senders. Consumers are believed to have no vested interest in recommending a product or brand, and their implied independence renders reviews to be more credible, and consequently, more useful than marketer-generated information (*see* BRAND COMMUNITY; ADVERTISING AND THE INTEGRATED MARKETING COMMUNICATIONS (IMC) PROCESS; SOCIAL NETWORKS; SOCIAL INFLUENCE).

Digital media not only increases consumers' control over brand communication but they also give advertisers many new possibilities. First, advertisers can (content) analyze the consumer-generated brand information and use it for targeting, customization, and product development. Second, advertisers can use consumer databases to personalize their messages by means of identifying the recipient (by name) and by matching the content of the message to known characteristics and interests of the recipient. Third, they can capitalize on the trend of consumers to actively process brand information by viral marketing (i.e., brand information that is passed along from person to person).

Traditional media. New developments in media can be found not only in digital media but also in traditional media. For example, newspapers have transformed and do look more like magazines with special interest sections, full color print, and lifestyle topics. Out-of-home media are now capable of interacting with the person in front of the billboard by using audio and movements.

The widespread penetration of DVRs and PVRs has resulted in a substantial reduction of live TV viewing time. Instead, a lot of viewing now is recorded viewing time which causes a problem for advertisers because of the risk that ads during recorded programs are skipped. On the other hand, TV offers new advertising possibilities including split screen advertising, EPG (Electronic Program Guide) advertising, and interactive ads. Television is becoming an advertising targeted, participatory, on-demand, nonlinear, two-way communication platform.

Other developments in traditional media have led to a blurring of the differences between media types. This process is called *media convergence*,

which makes it possible to receive the same content on different platforms. TV, for example, can not only be received on a television but also on a mobile phone and via the Internet, and is more selective due to cable and satellite offerings.

Other developments cause a transformation of the traditional media as well. Traditional lines between advertising and editorial content have blurred by presenting brand information in editorial articles and features. This integration of persuasive messages into originally nonsponsored content can be found in a variety of traditional and new media and is known under different names, such as product placement, brand placement, product integrations, advertorials, advergames, and sponsored media content (*see* PRODUCT PLACEMENT). Advertisers assume that entertainment contributes to attention for their messages. Moreover, sponsored media are more persuasive because audiences may not recognize its persuasive intent as a result of which cognitive defenses against persuasion, which are often automatically raised when encountering advertising, may not be activated. The worldwide increase of sponsoring of sport and music events is a related trend based on the same premises.

With respect to intertwining of advertising and entertainment, several theories in different disciplines (e.g., communication science, media psychology, health education) gradually begin to acknowledge that entertainment is processed in a unique way that differs from processing other communication forms. This unique processing, which has been named experiential, implicit, or heuristic processing in different fields, differs from other, more rational communication processing forms in that it is primary pleasure-oriented and characterized by a primacy of emotional reactions. These processes call for implicit measures of advertising effects, next to the traditional explicit measures.

Conclusion

To finalize, this article shows that traditional mass media such as television, radio, newspapers, magazines, and out-of-home are still very important advertising media, and that classical concepts such as reach, frequency, timing, and costs are still the central metrics for advertising buying and selling. However, new developments in advertising media are becoming increasingly significant and transform the media landscape. These developments can be characterized as more multimedia, targeted, customized, and personalized as well as integrating advertising and transaction. In today's media landscape, advertisers and consumers increasingly change roles in the communication process which places consumers more often in the driving seat.

Bibliography

Dijkstra, M., Buijtels, H., and Van Raaij, W.F. (2005) Separate and joint effects of medium type on consumer responses: a comparison of television, print, and the Internet. *Journal of Business Research*, **58** (3), 377–386.

Ephron, E. (2000). Media mix optimisers. Admap, March. Retrieved from http://www.warc.com.

FEN (Future Exploration Network) (2007) *Future of the Media Report 2007*, Future Exploration Network, Sydney.

Harvey, B. (1997) The expanded ARF Model; bridge to the accountable advertising future. *Journal of Advertising Research*, **37** (2), 11–20.

Heflin, D.T.A. and Haygood, R.C. (1985) Effects of scheduling on retention of advertising messages. *Journal of Advertising*, **14** (2), 41–64.

Jones, J.P. (1995) *When ad Works. New Proof that Advertising Triggers Sales*, Lexington Books, New York.

Kelly, L.D. and Juggenheimer, D.W. (2004) *Advertising Media Planning. A Brand Management Approach*, M.E. Sharpe, Inc., New York.

Kent, R.J. (1995) Competitive clutter in network television advertising: current levels and advertiser responses. *Journal of Advertising Research*, **35** (1), 49–57.

Krugman, H.E. (1965) The impact of television advertising: learning without involvement. *Public Opinion Quarterly*, **29**, 349–356.

Krugman, H.E. (1972) Why three exposures may be enough. *Journal of Advertising Research*, **12** (6), 11–14.

Ligthart, J. (1999) Planning op bereik vergroot effectiviteit tv-reclame (Planning on the basis of reach improves advertising effectiveness). *Adformatie*, **16**, 56.

McDonald, C. (1996) *Advertising Reach and Frequency: Maximising Advertising Results Through Effective Frequency*, 2nd edn, ANA/NTC, Chicago.

Moorman, M., Neijens, P.C., and Smit, E.G. (2005) The effects of program responses on the processing of commercials placed at various positions in the program and the block. *Journal of Advertising Research*, **45** (1), 49–59.

Naik, P.A. and Raman, K. (2003) Understanding the impact of synergy in multimedia communications. *Journal of Marketing Research*, **40** (4), 375–388.

Naples, M.J. (1979) *Effective Frequency: The Relationship Between Frequency and Advertising Effectiveness*, Association of National Advertisers, Inc., New York.

Norris, C.E. and Colman, A.M. (1992) Context effects on recall and recognition of magazine advertisements. *Journal of Advertising*, **21** (3), 37–46.

Park, D.-H., Lee, J., and Han, I. (2007) The effect of online consumer reviews on consumer purchasing intention: the moderating role of involvement. *International Journal of Electronic Commerce*, **11**, 125–148.

Rossiter, J.R. and Percy, L. (1998) *Advertising Communications and Promotion Management*, McGraw-Hill, Boston.

Smit, E.G. and Neijens, P.C. (2000) Segmentation based on affinity for advertising. *Journal of Advertising Research*, **40** (4), 35–43.

the role of signage in marketing: outdoor advertising, out-of-home media, and on-premise signs

Charles R. Taylor

Recent Developments in Outdoor Advertising and Out-of-Home Media

In contrast to on-premise signage, outdoor advertising has long been recognized as a traditional advertising medium. Although understudied, more attention has been paid to outdoor advertising. Outdoor advertising has not traditionally accounted for a large proportion of advertising budgets. However, overall expenditures have grown considerably in recent years, with annual expenditures in the United States alone having exceeded $5 billion in 2000 and every year since then (Outdoor Advertising Association of America, 2008).

In recent years, various types of out-of-home media, such as transit advertising (e.g., on trains and subways), kiosks, street furniture, bus shelters, advertising in airports, malls, health clubs, and other retail establishments, and others have also grown considerably. According to the Outdoor Advertising Association of America, out-of-home media, while accounting for just 2% of overall advertising expenditures, have been growing at a rapid pace (c.f. Wilson and Till, 2008).

As outlined by Taylor (1997), recent years have seen a shift toward a higher proportion of outdoor ads being for retail and service businesses. While there are still important uses of outdoor advertising for manufacturers, this trend makes sense given the primary advantages of outdoor advertising and is verified by the fact that all eight of the top outdoor advertising categories are retail and service businesses, including categories such as local services and amusements, hotels, retailers, media companies, restaurants, auto dealers, and financial services.

Studies by Lilley and DeFranco in 1995 verified that billboard users were predominately locally run businesses. A study of small towns found that 90% of billboard advertisers were locally owned (Lilley and DeFranco, 1995a) and that a large majority were small and medium sized businesses. A similar study of large cities such as Charlotte in North Carolina, and Indianapolis in Indiana found that 75% of those advertising on billboards were locally owned businesses (Lilley and DeFranco, 1995b).

There is considerable evidence that billboards and other out-of-home media provide unique advantages over other media to the types of firms who use the medium. Taylor and Franke (2003) surveyed billboard users and found that these users did not see other individual media as substitutes for billboards. Billboards were rated higher by their users compared to other media in terms of ability to (i) communicate information affordably, (ii) attract new customers, and (iii) increase sales (*see* ADVERTISING MESSAGE APPEALS). Moreover, most billboard users (75%) reported that they would lose sales if they did not have access to billboards, and the average estimated loss of sales among these users was 18.4%, and 14.2% for the overall sample.

As is well documented in marketing and advertising textbooks, billboards have several unique advantages relative to other advertising media. Textbook authors and academic researchers have identified a variety of distinctive characteristics of billboards and outdoor advertising (e.g., Belch and Belch, 1994; Kelley and Jugenheimer, 2004; Sissors and Baron, 2002; Taylor, 1997; Vanden Bergh and Katz, 1999). The advantages of billboards include (i) potential placement of the advertisement close to the point of sale, (ii) high frequency of exposure to regular commuters, (iii) high reach, (iv) 24-hour presence, (v) geographic flexibility for local advertisers, (vi) economic efficiency in terms of low production costs and low cost per thousand exposures, (vii) visual impact from advertisement size and message creativity, and (viii) brand awareness, among others. The disadvantages of billboards include (i) the need to limit the number of words in the message, (ii) short exposure to the advertisement, (iii) low demographic selectivity, and (iv) measurement problems. Drawing on prior discussion of the medium's advantages, Taylor, Franke, and Bang (2006) surveyed billboard users pertaining to the key reasons they continue to use billboards. A factor analysis revealed four key underlying dimensions that users associated with billboards:

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

1. *Visibility*: this factor pertains to billboards being easily seen, making a powerful impression, and being visible 24-hours a day.
2. *Media efficiency*: this pertains to billboards having high reach and frequency as well as being cost effective on a cost per thousand exposures basis.
3. *Local presence*: this reflects the ability of billboards to generate awareness/interest in close proximity to the location of a business, to help maintain a brand presence in a local market, and to provide a "last" close to the time of purchase decision.
4. *Tangible response*: the users reported that billboards are effective at generating traffic and helping to drive sales.

What is surprising, or at least interesting, about the advantages of billboards is how uniquely suited they are to certain retail and service business, and especially to many smaller businesses, who cannot afford other media. Taylor, Franke, and Bang (2006) theorized that traditional advantages of billboards had been heightened by outdoor advertising's ability to eliminate issues associated with selective perception and geographic advantages consistent with gravity models. In essence, they observed that billboards are simultaneously effective at cutting through an increasingly cluttered advertising environment and capitalizing on the localized nature of the medium. In essence, and consistent with Ephron's (1997) concept of recency planning, outdoor ads often appear in the right geographical location (*see* PRODUCT PLACEMENT) (close to the business) and at the right time (when the consumer is in a position to make a purchase decision). In this way, outdoor ads often provide a very unique advantage to certain retail and service businesses.

Taylor, Franke, and Bang (2006) also examined creative and executional factors associated with a successful billboard campaign. On the basis of a survey of managers, the following key factors were found:

1. *Name identification*: the billboard must be designed so that the name registers quickly and therefore the name is featured prominently.
2. *Location of the billboard*: it must be located at a good site.
3. *Readability*: the typeface should be legible, there should be a clear contrast, and strong colors should be used.
4. *Clarity*: the message should be communicated efficiently, there should be a simple background, and there should not be more than seven or eight words of text.
5. *Integrated marketing communications*: the billboard should reinforce what is communicated in other customer contacts.
6. *Visuals*: powerful visuals and effective illustrations have a positive impact.
7. *Creative*: clever slogans and novel copy increase effectiveness.
8. *Information*: the message should indicate benefits and uniqueness of product or service.

Recently, digital technology has added a new dimension to billboards. A typical digital billboard using LCD technology and rotating four to eight ads per minute can generate more revenue for the billboard company than a traditional billboard. Moreover, these billboards bring back access to prime locations to some advertisers who have not had access to them.

While there is debate in some municipalities as to whether, and to what extent, to allow digital billboards, public opinion polls conducted recently by Arbitron indicate that the general public sees significant advantages of billboards. A 2008 study of Cleveland consumers documents that digital billboards are very prone to be noticed and that a large majority of the public thought these billboards help communities communicate emergency information (Arbitron, 2008). The study also found that most of the Cleveland area residents surveyed thought that digital billboards are attractive, are a "cool way to advertise", and have "current and relevant information." A 2009 study of Los Angeles consumers by Arbitron (2009) affirmed that very similar results hold. Again, most consumers thought billboards help to communicate emergency information, are a "cool way to advertise," and have "current and relevant information." A majority of consumers also believed that digital billboards are attractive.

On-Premise Signs and their Use as Promotional Devices

An on-premise sign can be defined as (Taylor, Claus, and Claus, 2005, p. 1.6) follows: "A sign whose message and design relates to a business, profession, product, service, event, or other commercial activity sold, offered, or conducted on the same property where the sign is located. The structure and fact are designed (*see* PRODUCT DESIGN) in unison and are both integral parts of the sign's message. These signs may be owned or leased."

The service sector of the US economy accounts for about 80% of jobs (Winer, 2004). Considering their importance to most retail and service businesses, there has been very little study of the marketing functions of on-premise signs. However, a framework developed by Taylor, Claus, and Claus (2005) provides some insight into the primary functions of on-premise signs. In particular, on-premise signs can play at least four significant roles in marketing and advertising programs. These include

- communicating the location of the business;
- reinforcing advertisements and other marketing variables as part of integrated marketing communications;
- branding the site;
- enhancing store image.

Communicating the location of the business. As a result of performing these functions, on-premise signs contribute to the performance and sales of many businesses in terms of both traffic generation and other less tangible ways. For many retail and service businesses, a fundamental role played by the sign is communicating the location of the business to customers. An on-premise sign often allows the potential customer to learn where the business is located. Thus, the on-premise sign plays a key role for the business by "getting the word out" to the consumer that the business exists at a given location. For certain types of businesses, such as those that rely on impulse purchases, or serve immediate needs, the sign may often induce an immediate response that leads to a sale.

Signs are also important way finding devices. For consumers passing by the business, high visibility of the sign is needed to help the consumer identify where the store is (Berman and Evans 2007). Thus, on-premise signs help consumers find stores and help businesses by generating traffic.

Reinforcing advertisements and other marketing variables as part of integrated marketing communications. Clearly, on-premise signs function as part of integrated marketing communications. Marketing communications must now be viewed holistically (e.g., Schultz, Tannenbaum, and Lauterborn, 1995), and the idea of integrated marketing communications emphasizes that every point of contact a company has with the consumer must be coordinated in order to maximize the positive impact on the brand (Duncan, 2005). The way in which consumers perceive the brand will result from the synthesis of messages they receive at contact points with the business (Belch and Belch, 2004). Clearly, on-premise signs are an important point of contact that must be viewed as part of a company's integrated marketing communications strategy. An attractive sign can contribute to positive consumer perceptions of the company. By displaying a logo or brief message, it can also serve to reinforce other messages sent by the marketer, thereby capitalizing on the sign as part of an integrated marketing communications program.

Branding the site. It has been well established that today's marketers need to think in terms of building brand equity (*see* PERCEPTION OF BRAND EQUITY) (Keller, 2000). Building brand equity can help a firm increase customer loyalty, decrease price sensitivity, and increase awareness of the company and its brands. (Yoo, Donthu, and Lee, 2000).

For retail and service businesses, the on-premise sign plays a unique and critical role by bringing branding to the site of the business. Trade dress, including trademarks and other aspects of trade dress, can be used both to enhance immediate recall of the business when passing by and to build positive associations with the business in the consumer's mind. Exposure to trade dress via the on-premise sign can be critical in shaping a business' image. Additionally, the need to reinforce the brand at the point of purchase for retail and service

businesses is fundamental to maximizing brand equity (Taylor, Claus, and Claus, 2005). For these reasons, we propose that on-premise sign helps businesses to build brand equity with customers.

Enhancing the image of the store or business. Scholars have long maintained that most brick and mortar retailers must create and reinforce a store image in order to ensure success. Signage can play an important role in communicating the store's image to consumers (Berman and Evans, 2007). The on-premise sign can communicate a high-end image by using expensive signage in conjunction with an elegant storefront. Regardless of the image of the retailer, the on-premise sign should demonstrate consistency with it. It is well documented that retail stores and service businesses need to project a well-defined, distinctive image to consumers in order to be successful.

Collectively, these four marketing and advertising functions of on-premise signs make their effective use important to the promotional programs of retail and service businesses. With the increasing use of digital technology and light emitting diodes (LED) technology, there are now more options for using the on-premise sign- to communicate additional information to consumers. Messages promoting the product or services offered can be transmitted digitally, allowing on-site communication with consumers.

Selecting the right on-premise sign for the business. There are several factors that a business should consider in selecting an appropriate on-premise sign in order for it to play the appropriate promotional role for the business (*see* PRODUCT DESIGN). First, the sign must be of sufficient size for consumers to be able to notice and easily read it. Factoring in that most states allow drivers with 20/40 vision to obtain licensing, a good guideline is that one inch of letter height is needed for every 25 feet of distance from which the business would expect the sign to be readable.

A second factor in choosing the right on-premise sign is ensuring that it is visually conspicuous. In addition to the need to be sufficiently large to be read, an effective sign needs to stand out from its visual surroundings. When signs are placed in urban environments, they compete with other signs and landscape features in order to be noticed. In order to be conspicuous, the sign must contrast with its background and be easily detectable in the context of the rest of the visual environment.

The height and placement of the sign are also important in allowing it to perform its communications functions. The appropriate height must factor in the speed at which traffic travels and the speed limit. On a four lane highway with a 55 mile per hour speed limit, for example, a sign should be 55 feet high (for more information, see Taylor, Claus, and Claus, 2005 on page 8.5). Placement refers to making sure the sign is placed at a location that will allow a motorist to respond in time upon seeing the sign. This factor should be considered in site selection.

For businesses that operate at night or wish to promote themselves at night, proper illumination of the sign is a fifth important factor to consider in obtaining an on-premise sign. Lighting very much impacts the effectiveness of the sign.

Perhaps most important from a promotional standpoint is the need for the sign to communicate something about the business. An effective sign will communicate something to people about what they will find on the inside of the business. Here, the design, color, and content of the sign must be carefully considered. Logos and trade dress can often be used effectively, and other messages about the business can sometimes be communicated either through lettering or, more recently, through new technologies.

Summary and Conclusion

Outdoor advertising and the marketing functions of on-premise signs have been understudied topics for many years. However, given their importance to business and a more cluttered advertising and promotional environment, it is likely that more attention will be paid to signage by academics, going forward. Research has documented that outdoor advertising is important to the very viability of many businesses, most of which are small or medium-sized local businesses. Meanwhile, larger national advertisers can use out-of-home media effectively as part of

an overall integrated marketing communications strategy.

On-premise signs have long been very important to businesses but have been the subject of very few academic studies in spite of the role they play for many businesses. Like billboards, they can be especially important to small and medium-sized businesses in terms of their viability. Going forward, more research on the marketing roles of on-premise signs is needed.

Bibliography

Arbitron (2008) *Arbitron Digital Billboard Report: Cleveland Case Study*, Arbitron, New York.

Arbitron (2009) *Arbitron Outdoor Billboard Report: Los Angeles County Case Study*, Arbitron, New York.

Belch, G.E. and Belch, M.A. (1994) *Introduction to Advertising and Promotion: An Integrated Marketing Communications Perspective*, Richard D. Irwin, Chicago.

Belch, G.E. and Belch, M.A. (2004) *Advertising and Promotion: An Integrated Marketing Communications Perspective*, McGraw Hill. Irwin, Burr Ridge.

Berman, B. and Evans, J. (2007) *Retail Management*, 10th edn, MacMillan, New York.

Duncan, T.R. (2005) IMC in industry: more talk than walk. *Journal of Advertising*, **34** (4), 5 –6.

Ephron, E. (1997) Recency planning. *Journal of Advertising Research*, **37** (4), 61–66.

Keller, K.L. (2000) The brand report card. *Harvard Business Review*, **78**, 3–10.

Kelley, L.D. and Jugenheimer, D.W. (2004) *Advertising Media Planning: A Brand Management Approach*, M.E. Sharpe, Armonk.

Lilley, W. and DeFranco, L.J. (1995a) *The Economic Impact of Outdoor Advertising in Rural, Small-Town America*. Report prepared by Incontext, Inc., 1615 L. St., NW, Washington, DC, p. 20036.

Lilley, W. and DeFranco, L.J. (1995b) The Economic Impact of Outdoor Advertising on Urban, Big-City America. Report prepared by Incontext, Inc., 1615 L. St., NW, Washington, DC, p. 20036.

Outdoor Advertising Association of America (2008) Q1 Revenue Growth Out of Home Revenue Grows 3.0%, http://www.oaaa.org/marketingresources/revenue/q1revenuegrowth.aspx (accessed 28 September 2009).

Reilly, W.J. (1931) *The Law of Retail Gravitation*, Knickerbocker Press, New York.

Schultz, D.E., Tannenbaum, S.I., and Lauterborn, R.F. (1995) The New Marketing Paradigm: Integrated Marketing Communications, NTC, Chicago.

Sissors, J.Z. and Baron, R.B. (2002) *Advertising Media Planning*, 6th edn, McGraw-Hill, New York.

Taylor, C.R. (1997) A technology whose time has come or the same old litter on a stick? An analysis of changeable message billboards. *Journal of Public Policy and Marketing*, **16** (2), 179–186.

Taylor, C.R., Claus, T.A., and Claus, S.L. (2005) *On-Premise Signs as Storefront Marketing Devices*, U.S. Small Business Administration, Washington, DC.

Taylor, C.R. and Franke, G.R. (2003) Business perceptions of the role of billboards in the U.S. Economy. *Journal of Advertising Research*, **43**, 150–161.

Taylor, C.R., Franke, G.R., and Bang, H. (2006) Use and effectiveness of billboards: perspectives from selective-perception theory and retail-gravity models. *Journal of Advertising*, **35**, 21–34.

Vanden Bergh, B.G. and Katz, H. (1999) *Advertising Principles: Choice, Challenge, Change*, NTC Press, Lincolnwood.

Wilson, R.T. and Till, B.D. (2008) Airport advertising effectiveness: an exploratory field study. *Journal of Advertising*, **37** (1), 59–72.

Winer, R.S. (2004) *Marketing Management*, 2nd edn, Prentice Hall, Upper Saddle River.

Yoo, B., Donthu, N., and Lee, S. (2000) An examination of selected marketing mix elements and brand equity. *Journal of the Academy of Marketing Science*, **28** (2), 195–211.

trade promotions: issues and findings

Naveen Donthu and Amit Poddar

Trade promotions are any promotions that are targeted toward the trade (supply chain). They come in many forms, such as cash discounts, off-invoice, scan backs, quantity discounts, and so forth. They are generally classified as push promotions as the goal of the promotion is to induce the trade to push the promoted product. According to some estimates, trade promotions overall command 52% of the total money spent on advertising and promotion (1999), whereas the rest of the money is shared between advertising and consumer promotions.

Our goal in this article is to bring together important findings in the trade promotion literature and summarize what we know about trade promotions. We have classified the findings into eight substantive questions.

1. Why do firms offer trade promotions?
2. What is the value of trade promotions?
3. What are the factors that lead to trade promotion success?
4. Can trade promotions be explained as power games?
5. What are the costs and inefficiencies of trade promotions?
6. Can trade promotions be improved?
7. What are the main problems with trade promotions?
8. Are slotting fees good for firms?

Why do Firms Offer Trade Promotions?

Researchers and practitioners know that trade promotions do not build long-term consumer franchise. They also know that retailers often do not pass on trade deals to consumers. Then the question is why firms offer trade deals to retailers instead of just sticking with consumer promotions or advertising? Existing research offers three basic theories to explain why firms offer trade promotions: inventory holding cost theory, weak customer equity theory, and colluding marketer theory.

Inventory holding cost theory. Trade promotion logic describes that retailers who receive the promotion pass on some of the money to consumers as price promotions, which encourages trial. This seems to be a reasonable explanation for new-product introductions, but this does not explain why trade promotions take place in mature markets. It can be argued that retailers use price promotions to shift inventory holding costs to the consumers, because consumers have lower holding costs. Cui, Raju, and Zhang (2008) argue that big retailers may also have lower holding costs compared to manufacturers (or vendors), in which case it makes sense for the manufacturer to shift some of its own holding costs to the retailer through trade deals. The effect of this inventory holding cost transfer benefits both the retailer and the vendor, and the inventory cost differential allows vendors to price discriminate using trade promotions.

Weak customer equity theory. Although the inventory cost theory makes intuitive sense, trade deals also might result from the customer equity of the firms. According to Raju, Srinivasan, and Lal (1990), in a competitive market with different levels of brand equity, strong brands know they can lure a weaker brand's customers away with price discounts. To defend its turf, the weaker brand uses promotions to keep its customers, and the result is that all brands promote. This line of research also indicates that weaker brands tend to gain more from temporary price discounts. However, if all competing brands have high brand loyalty, pure equilibrium exists, and no one promotes. Therefore, trade promotions are inevitable whenever there are weaker brands in the market.

Colluding marketer theory. Another interesting explanation of trade discounts comes from Lal (1990), who suggests that many brands promote as a form of implicit collusion to prevent encroachment by private brands. In a study of the beverage market, Lal finds that national brands (e.g., Coca-Cola, Pepsi) fight competition from local brands by reducing their prices in alternate periods in an infinite horizon game.

What is the Value of Trade Promotions?

It is well accepted that trade promotions lead to increased sales in the short run. Researchers

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

and practitioners often wonder if increased sales translate into increased value captured by the firm in the form of incremental revenue minus the cost of the promotion.

When manufacturers provide a trade deal, they incur a cost, equal to the income transferred from them to the retailers. However, this income transfer is mitigated by the increased sales of their products. As long as the increased sales are greater than the increased costs, the process benefits the manufacturer, and the value is positive. The manufacturer faces two caveats though: It must know its baseline sales before it can claim an increase in sales, and the trade promotions should not just prompt sales that the manufacturer would have made anyway in the future.

The impact on retailers' profits is more complicated. Retailers benefit from trade promotions in many ways. First, they might buy at discounted prices and sell at normal prices. Second, trade promotions might allow them to sell more of the promoted product when they pass on the savings. These increased sales lead to higher inventory turnover. The discounted price may also allow the retailer to claim price leadership, which can lead to more sales in different product categories. However, retailers also incur costs that are difficult to capture, such as the costs attached to forward buying, in that buying more than required increases inventory holdings, which reduces profits. When retailers promote some products, they also might suffer reduced sales of nonpromoted products, which could have higher margins. Finally, the reduction in sales of other products could lead to higher inventory costs for the nonpromoted products, which adds to the retailers' losses. Existing research has only been partially successful in calculating the actual worth of trade deals. The jury is still out regarding the impact of trade deals on profitability.

Value of trade deals for the manufacturer. Researchers have long speculated that trade promotions may be value losers, causing them to investigate both the long- and short-term impacts of trade promotions. Manufacturers blame retailers for taking advantage of trade promotions but not passing on the benefits to the ultimate consumers. It is estimated that approximately 30% of the money spent on trade promotions goes straight to the bottom line of the retailers (Kasulis *et al.*, 1999). Blattberg and Levin (1987) attempt to model how retailers might behave when offered trade promotions and which promotions would be best for them. They note that if promotions do not increase consumer sales, they merely shift the timing of the retailer's purchase. Empirically, they find that trade promotions do not pay for manufacturers. Overall, researchers seem to agree that trade deals look like a losing proposition for manufacturers.

Value of trade deals for the retailer. The impact of trade promotions on retailer profit may not be clear because of various factors such as forward buying, reduced profitability of promoted items, reduced sales of nonpromoted items because of consumers buying promoted items, and the possibility that consumers drawn to stores could end up buying other products that actually increase overall sales.

Trade promotions may favor retailers that forward buy but not manufacturers when the retailer's promotions reduce the brand value. Similarly, promotions may be favorable to manufacturers and unfavorable to retailers when increased price cutting occurs at the retailer level to prevent undercutting by competitors. If all the costs of trade promotions, including advertising, display setup, and display rent, are charged to the promotions, their value might become negative.

Recent research indicates a negative net profit impact of trade promotions. Ailawadi *et al.* (2006) find that trade promotions provide an incremental sales lift of 45%, but the profit impact is negative because the products that create higher promotional bumps earn lower margins overall. This problem gets aggravated when all promoted products earn the lower margin, not only the incremental units. They also find that despite sales increases, more than half the promotions are not profitable for the retailer.

What are the Factors That Lead to Trade Promotion Success?

Because not all trade promotions succeed in creating value, researchers have attempted to

pinpoint the factors required for a successful promotion. Overall, managers realize that considering the amount of money at stake, it makes sense to develop a trade promotion that works.

Management-based success criteria. Hardy (1986) contends that to determine success, the objectives of the promotion have to be specified first. He lists the following main objectives of trade promotions:

1. Achieve short-term volume.
2. Achieve long-term market share.
3. Build trade inventories.
4. Increase consumer trials.

He also proposes that these objectives are dependent on the promotion period, promotion cost, trade support, presence of competitive promotions, level of incentives, and presence of consumer promotions alongside trade promotions. His empirical tests reveal that achieving trade support is the most critical factor for achieving the objectives of the firm. Hardy (1986) also finds that in the view of managers, trade promotions succeed only when high incentives, good trade support, good sales force support, and an absence of competitive activity exist. The reasons for unsuccessful trade promotions include builtup inventory from previous deals, insufficient incentives, competitive promotions, and a lack of trade support.

Deal-based success criteria. Walters's (1989) empirical investigation attempts to determine the level of retailer trade support, thus delving deeper into Hardy's (1986) assertion that trade support represents the most critical success factor in trade promotions. Walters uncovers empirical evidence that suggests that economic incentives contained in deals, such as advertising support, price reductions, and product displays, significantly affect retailers' support of trade deals. He also finds that the time since the last promotion in the same category influences retailer support, though he does not find any significant impact of product-related factors (e.g., store sales rank of product category and product sales rank in category), manufacturer's consumer promotions, or the price elasticity of the product on trade support. The finding that consumer promotions have no impact on trade support seems surprising, because conventional wisdom indicates retailers should be more willing to provide trade support when trade deals accompany consumer promotions. It also contradicts Hardy's (1986) findings that consumer promotions, run simultaneously with trade promotions, affect the success of trade promotions.

Can Trade Promotions be Explained as a Power Game?

Power plays an important role in the relationship between the retailer and the vendor. Because trade promotions entail income transfers, power plays an important role in determining the extent and the form of transfer. Understanding the power conundrum and its role remains an important research problem, for which researchers use game theory to predict why trade promotions take place and explain the different kinds of trade deals that may appear in different power asymmetry situations.

Game theoretic explanation of trade promotions. Researchers have tried to study trade promotions as a sort of prisoner's dilemma, which would imply one best strategy for a firm, irrespective of what its competitor does. Rao, Arjunji, and Murthi (1995) model trade promotions as a series of competitive games and hypothesize about the actions of the parties, depending on a series of promotion outcomes. For example, they predict that if a promotion is profitable, regardless of what the competition does, it represents a prisoner's dilemma game, and both parties end up promoting. However, if the promotion is profitable only if the competitor does not promote, the game type is a "battle of sexes," and it is not possible to go without promotions over a long period of time. They demonstrate empirically that promotion activity seems independent of competitor actions; that is, competitors do not necessarily choose their promotion actions after they take into account the actions of their competitors.

Power asymmetry and impact on different forms of trade deals. Kasulis *et al.* (1999) also address trade promotions as the result of a power game

between the manufacturer and the retailer. Considering the relative power between the two parties, they hypothesize that the use of different forms of trade promotions has different objectives. Their conceptual framework consists of a two by two matrix (high versus low) of retailer and manufacturer power, such that when the retailer dominates, they anticipate promotions that shift channel profits from the manufacturer to the retailer (e.g., bill backs, slotting fees, and inventory financing). When the manufacturer is more powerful than retailers, it sees little point in offering trade deals and instead focuses on consumer promotions that increase customer loyalty. In the symmetric case when both the retailer and the supplier are strong, Kasulis and colleagues propose a higher incidence of promotions such as co-op advertising, display advertising, and calendar marketing agreements, whereas in a weakly symmetric case, both parties enter survival mode and likely use promotions that result in temporary price cuts passed on to consumers.

G'omez, Maratou, and Just (2006) find that as the power of the manufacturer increases, the amount of funds allocated to trade deals declines, whereas greater retailer bargaining power leads to more focus on deals such as off-invoice agreements. Retail companies with larger shares of private labels and greater annual sales can increase the allocation of promotional funds toward off-invoice agreements and away from performance-based trade promotions, such as scan backs and bill backs (G'omez, Rao, and Mclaughlin, 2007). Although manufacturers may have greater autonomy in determining trade promotion budgets, retailers have a greater say in the allocation of the funds. The authors find only mild support for the notion that the retailer and vendor decide on the overall budget and allocation pattern in a joint negotiation process. Similarly Ailawadi and Harlam (2009), considering the role of power with respect to retail pass-through, note that higher manufacturer power increases pass-through, but higher retail power is associated with lower pass-through.

What are the Costs and Inefficiencies of Trade Promotions?

Researchers worry that trade promotions might add costs that get passed on to the ultimate consumer. Researchers thus have tried to calculate the cost that is added to products due to trade promotion.

Calculating the dollar cost of trade deals. Buzzell, Quelch, and Salmon (1990) calculate that trade promotions add 0.5–1.1% to total retail prices. To calculate these costs, they note that trade promotions (especially price promotions) have distinct impacts on retailer behavior. Trade promotions lead to the phenomenon of forward buying, such that retailers take advantage of the lower prices to forward buy for later sales at normal prices. Retailers have a motivation to forward buy until the savings from the lower prices equal their holding costs. They calculate the holding costs to be approximately 30%, which include handling, storage, and capital charges. Their article also highlights two main negative effects of trade promotions: (i) they lead to distrust between the manufacturer and distributor, which could prompt higher transaction costs; and (ii) forward buying leads to wasteful storage and diversion expenses that help no one in particular.

The cost burden of trade promotions is not borne by everyone in equal measure. People classified as "cherry pickers" (bargain hunters) get a lower price, whereas normal, brand-loyal consumers end up paying higher prices. Therefore, loyal consumers subsidize nonloyal consumers, which is bad for the manufacturer in the long run.

Dangers of going without trade deals. Prior to 1995, research held that forward buying adds inefficiency to the channel system, ensuring that the manufacturer is always worse off. Trade deals thus had to be unprofitable, because the retailers would forward buy and keep the promotion to themselves. Lal, Little, and Villas-Boas (1996) challenge this deeply held belief by modeling the behavior of manufacturers, retailers, and consumers. They find that in equilibrium, manufacturers benefit when they allow forward buying compared with when they do not. The intuition behind

their surprising result states that allowing retailers to forward buy reduces the intensity of competition (between manufacturers), which helps the manufacturer. Moreover, forward buying decreases income for the manufacturer, but manufacturers still achieve a better position than they could have with the prices that would have resulted without forward buying and with a higher intensity of competition. Lal, Little, and Villas-Boas (1996), however, acknowledge that forward buying has a negative impact on total channel costs. It creates serious logistical dysfunctions, leading to the excess storage of inventory, which creates inventory storage costs.

Can Trade Promotions be Improved?

Considering that the current way of doing trade promotions creates inefficiencies in the supply chain, researchers have made considerable efforts to suggest ways in which trade promotion inefficiencies might be removed. Some of the most common suggestions include every day low purchase price (EDLPP), scan backs, and electronic forward buys. Suggestions have also been made to improve the retail pass-through, which is likely to improve the long-term benefits of trade promotions.

Evaluation of everyday low purchase price method. Buzzell, Quelch, and Salmon (1990) may have been the first researchers to suggest a policy for EDLPP to reduce the costs of running trade promotions. They argue that the retailer should purchase on an as-needed basis and receive a weighted average price that reflects both the deal price and the promoted price. There are three benefits of EDLPP. First, it prevents inventory buildup for both manufacturers and retailers; second, it reduces selling and administration expenses because retailers would spend less time negotiating trade deals; and third, it leads to a more collaborative relationship between the retailer and the manufacturer, in which both are freed from the zero sum game in which one party wins only at the expense of the other. They also posit that following an EDLPP approach leads to more pass-through, which lowers the prices for consumers.

Neslin, Powell, and Stones (1995) also mention decreasing pass-through, greater promotion intensity, and increased retailer warehousing ability (which allows for forward buying) as reasons for an EDLPP strategy by manufacturers. However, they warn that consumer factors may prevent a wholesale move to EDLPP. Consumer responses to deals remain intact, and media advertising appears to be losing effectiveness, which have opposing effects on the adoption of EDLPP. Further speculation indicates that promotional elasticities might far exceed price elasticities, in which case the EDLPP strategy becomes questionable. Mangers may prefer to keep a higher shelf price and then offer discounts to generate more sales and profits.

Although several authors praise EDLPP as a solution to deal-to-deal forward buying, Ailawadi, Farris, and Shames (1999) claim that it may be strong medicine that causes its own side effects. They decry the tendency by manufacturers to blame trade promotions, stating that in the absence of trade promotions, manufacturers would have to charge much less than their current list prices, in agreement with Lal, Little, and Villas-Boas (1996). They adopt a different view of how to improve trade promotions; the main problem with trade discounts, as they see it, is that they are linked to the quantity bought, which encourages forward buying and creates problems with pass-through. They instead suggest that the goal of the manufacturer should be to design trade promotions to increase total channel profit. With their analytical model they demonstrate that linking promotional allowances to the list price charged by the retailer increases total channel profit. The intuition behind their thinking is that if the retailer receives encouragement to charge a lower price (because the total allowance depends on the price charged to the consumer), total base demand increases, which should help both the retailer and the manufacturer.

Improving and increasing retail pass-through. Another problem that plagues the effective implementation of trade promotion is retail pass-through, which occurs in response to "retailer opportunism." Kumar, Rajiv, and Jeuland (2001) state that the manufacturer wants the retailer to pass on the trade promotion money to the consumer as reduced prices, but

the retailer does not because of its information asymmetry. The consumer has no way of knowing when the manufacturer has provided a trade promotion. If the retailer never passes on a trade promotion, the consumer might become suspicious and shop elsewhere, so the retailer occasionally charges a lower price in association with a trade promotion but charges a normal price on other occasions. They suggest and demonstrate that when manufacturers advertise their ongoing trade promotions directly to consumers, it increases retail pass-through and thus reduces retailer opportunism.

Few empirical articles review actual retail pass-through, though a recent study (Ailawadi and Harlam, 2009) suggests that retail pass-throughs for individual manufacturers reach approximately 75% and that retailers actually spend more than 100% of the funds they receive as trade promotions. However, the retailers shift their trade promotion dollars from individual vendors to private labels and categories with higher sales and higher market share. Thus, retailers appear to view trade promotions as a pool of money they can spend in whatever way provides them with the greatest benefit.

Evaluating performance-based trade deals – the scan back. The call to eliminate forward buying has led to the creation and advocacy of a new type of trade promotion, the "scan back," which has become increasingly popular among manufacturers. With a scan back, the retailer receives promotion money only when it sells the product to the final consumer (tracked by scanner data during the promotion period), not when it buys the product. This approach effectively prevents any forward buying, because if the retailer buys extra inventory that it cannot sell during the promotion period, it loses promotion money. Recommendations of the scan back scheme promote it as a panacea for trade promotion ills. However, Dreze and Bell (2003), using an analytical model, prove that when the terms of the trade are identical (i.e., base size, deal size, and deal duration), manufacturers always prefer scan backs and retailers always prefer off-invoice deals. Because scan backs have benefits, in terms of reducing forward buying (which adds cost to the overall channel system), they suggest that the only way manufacturers can convince retailers to adopt scan backs is to compensate them for the loss of profit opportunities. To prove that such a system is possible, Dreze and Bell devise a modified scan back (which they call a mimic scan back) that leaves the retailer weakly better off and the manufacturer strictly better off – a sort of win–win solution for both parties. They propose that to create a mimic scan back, the manufacturer must provide a smaller deal for a longer period of time and lower the base price of the product. They also find empirical evidence that scan backs lead to more retail pass-throughs and lower retail prices, which in turn result in greater sales.

Most of the research which suggests ways to improve trade promotion efficiency considers the problem only from the manufacturer perspective. Suggestions, such as scan backs, are very beneficial from the point of view of the manufacturer but would never work unless retailers are also benefited from them. Similarly, though EDLPP might look very attractive for the manufacturer, it is not very attractive to the retailer. The failure of these alternate systems to catch on suggests that retailers are not buying into systems that fail to demonstrate obvious benefits for them.

Identifying Problems with Trade Promotions

Managers and academics have long believed that trade promotions create more problems than they solve. They specifically are thought to have a strong effect on cooperation between business partners. The key finding is that because trade promotions are looked at as zero sum games, retailers and vendors end up competing with each other rather than cooperating.

Effect on competition and consumer behavior. Trade promotions create adversarial relationships between partners and also create competitive retaliation between competitors. The effect on the brand being promoted is also substantial. Trade deals end up stealing funds from advertising, which devalue the brand image/consumer franchise in the long run. Trade deals can have negative effects on the performance of manufacturer in the short and long run. First, trade deals

can increase the price of products in the marketplace, leading to reduced demand. Second, the increasing market price may prompt the retailer to enter the market with its own private brands, in which case the manufacturer has created its own competitors. Third, as a result of high inter-retailer competition, retailers may become more aggressive in demanding trade deals, which reduces trust between the retailer and the manufacturer. Fourth, high fluctuations in the market price may cause consumers to become more price conscious and turn them into constant deal seekers. Fifth, as manufacturers dedicate more and more money to trade promotions, their ability to create brand differentiation decreases over time. This development has a long-term impact on the ability to charge consumers a premium for the brand.

Are Slotting Fees Good for Firms?

No discussion about trade promotions would be complete without mentioning the current controversy about the appropriateness of slotting fees. Retailers and wholesalers frequently require manufacturers to pay some sort of fees before they agree to stock a new product. Although a form of trade promotion, the impact of these slotting fees does not apply to mature products. According to Sudhir and Rao (2006), slotting fees vary from $1.4 to 2 million for a national-level introduction of a single SKU. The ongoing controversy about slotting fees persists because researchers cannot agree about the role they play in manufacturer–retailer relationships. There are two main schools of thought with diametrically opposing views, namely, the efficiency effects and the market power effects of slotting fees.

Efficiency effects of slotting fees. The first school of thought, the "efficiency school," states that slotting fees actually increase efficiency in the system by

1. providing a signaling mechanism for manufacturers to advertise product quality.
2. sharing risks between the retailer and the manufacturer. Because of information asymmetry, the retailer likely knows less than the manufacturer about the probability of success of the product. Therefore, it suffers a disproportionate amount of risk for a product introduction. Slotting fees help maintain the balance by shifting the risk from the retailer to the manufacturer.
3. aiding in the efficient allocation of shelf space. Slotting allowances help the retailer make efficient use of its scarce shelf space. In effect slotting allowances help totally new products attain shelf space, whereas normally, they would have been rejected because they were untested in terms of marketplace performance.
4. increasing competition and thus reducing total retail prices.

The efficiency school of thought largely is favored by retailers that want slotting fees to continue.

Market power effects of slotting fees. The other school of thought, called "*market power*," argues that slotting fees actually are harmful and damage competition and overall consumer behavior by

1. allowing retailers to use their market power to demand and obtain fees. Retailers thus can demand more fees from smaller manufacturers.
2. undermining channel relationships, because manufacturers become bitter about being made to pay fees to get their products to the market.
3. providing a mechanism for price discrimination. When different manufacturers pay differential fees, the costs increase disproportionately.
4. introducing unfair competition for certain manufacturers that cannot pay the slotting fees and thus just quit the market. This is especially true for small startups who often cannot afford the fees to get on the shelves.
5. harming the consumer, because the fees ultimately are passed on to consumers in the form of higher list prices.

Bloom, Gundlach, and Cannon (2000) use a survey methodology to discover that retailers and manufacturers do not agree about slotting fees constituting a signaling mechanism, though they concur that the fees shift risk from the retailer to the manufacturer. Furthermore, both

retailers and vendors agree that slotting fees benefit large manufacturers and lead to higher prices. Overall, their research supports the power school of thought, though a recent article by Sudhir and Rao (2006) challenges their findings and suggests that the efficiency theories actually are right.

Conclusion

The industry already spends more on trade promotions than on any other marketing activity, yet we also know the least about it, which means that there is a significant opportunity for understanding and obtaining efficiency gains. Since the gains from trade promotion directly affect the expense side of profit-and-loss statements, companies could realize massive benefits from any improvements to trade promotions.

Bibliography

Ailawadi, K., Farris, P., and Shames, E. (1999) Trade promotion: essential to selling through resellers. *Sloan Management Review*, **41** (1), 83.

Ailawadi, K. and Harlam, B.A. (2009) Retailer promotion pass-through: a measure, its magnitude, and its determinants. *Marketing Science*, **28** (4), 782–791.

Ailawadi, K., Harlam, B.A., Cesar, J., and Trounce, D. (2006) Promotion profitability for a retailer: the role of promotion, brand, category and store characteristics. *Journal of Marketing Research (JMR)*, **43**, 518–535.

Blattberg, R.C. and Levin, A. (1987) Modelling the effectiveness and profitability of trade promotions. *Marketing Science*, **6** (2), 124.

Bloom, P.N., Gundlach, G.T., Cannon, J.P. (2000), Slotting allowances and fees: schools of thought and views of practicing managers, *Journal of Marketing*, Vol. 64 No. April, 92–108.

Buzzell, R.D., Quelch, J.A., and Salmon, W.J. (1990) The costly bargain of trade promotion. *Harvard Business Review*, **68** (2), 141.

Cui, T.H., Raju, J.S., and Zhang, Z.J. (2008) A price discrimination model of trade promotions. *Marketing Science*, **27** (5), 779–795.

Dreze, X. and Bell, D.R. (2003) Creating win-win trade promotions: theory and empirical analysis of scan-back trade deals. *Marketing Science*, **22** (1), 16.

G'omez, M.I., Maratou, L.M., and Just, D.R. (2006) Factors affecting the allocation of trade promotions in the u.s. food distribution system. *Review of Agricultural Economics*, **29** (1), 119–140.

G'omez, M.I., Rao, V.R., and Mclaughlin, E.W. (2007) Empirical analysis of budget and allocation of trade promotions in U.S. supermarket industry. *Journal of Marketing Research (JMR)*, **44** (3), p. 410–424.

Hardy, K.G. (1986) Key success factors for manufacturers' sales promotions in package goods. *Journal of Marketing*, **50** (3). 13–23.

Kasulis, J.J., Morgan, F.W., Griffith, D.E., and Kenderdine, J.M. (1999) Managing trade promotions in the context of market power. *Journal of the Academy of Marketing Science*, **27** (3), 320–332.

Kumar, N., Rajiv, S., and Jeuland, A. (2001) Effectiveness of trade promotions: analyzing the determinants of retail pass through. *Marketing Science*, **20** (4), 382.

Lal, R. (1990) Price promotions: limiting competitive encroachment. *Marketing Science*, **9** (3), 247–262.

Lal, R., Little, J.D.C., and Villas-Boas, J.M. (1996) A theory of forward buying, merchandising, and trade deals. *Marketing Science*, **15** (1), 21.

Neslin, S.A., Powell, S.G., and Stones, L.S. (1995) The effects of retailer and consumer response on optimal manufacturer advertising and trade promotion strategies. *Management Science*, **41** (5), 749.

Raju, J.S., Srinivasan, V., and Lal, R. (1990) The effects of brand loyalty on competitive price promotional strategies. *Management Science*, **36** (3), 276.

Rao, R.C., Arjunji, R.V., and Murthi, B.P.S. (1995) Game theory and empirical generalizations concerning competitive promotions. *Marketing Science*, **14** (3), G89.

Sudhir, K. and Rao, V.R. (2006) Do slotting allowances enhance efficiency – or hinder competition? *Journal of Marketing Research (JMR)*, **43** (2), 137–155.

Walters, R.G. (1989) An empirical investigation into retailer response to manufacturer trade promotions. *Journal of Retailing*, **65** (2), 253.

business-to-business media selection

Kyle Krueger and Lawrence Soley

Advertising media selection for business-to-business (B2B) marketers is a process similar to that used with consumer advertising (*see* ADVERTISING MEDIA SELECTION). A media planner working for a B2B advertiser is concerned with selecting the proper media (e.g., trade publications) and vehicles (e.g., *Builder Magazine*) to reach the target market (e.g., remodeling contractors). The selection method consists of making qualitative decisions, such as choosing vehicles that provide a suitable editorial setting for the product or service being advertised, and quantitative judgments, such as calculating the vehicle's relative cost efficiency. A B2B media planner looks to efficiently select media that achieve the advertiser's communication objectives, such as building brand awareness, and marketing objectives, such as customer retention and loyalty (*see* ADVERTISING EFFECTIVENESS).

Historically, the media type preferred by the B2B media planner has been trade publications. Even in the age of Internet advertising, trade publications still account for 13% of B2B media spending. Business executives spend an average of 2 hours per week reading industry-related magazines (Sorce and Dewitz, 2007). Other preferred media channels include direct mail, trade shows, affiliate marketing, and digital media, including email, internet display ads, CDs, and search engine marketing. With trade publications, these media types account for 60% of B2B advertising expenditures (Sorce and Dewitz, 2007). One reason media spending for digital media is higher than for print publications is that most business publications produce online editions, blogs, and forums, and distribute e-newsletters. Many trade publications also lease their subscriber lists for direct mail.

When evaluating the cost efficiency of a trade publication, a media planner will usually use circulation figures, rather than readership, as is done with consumer publications. This is because there are no syndicated services measuring trade publication audiences, as there are for consumer magazines and newspapers. In contrast with business publications, Mediamark Research, Inc. (MRI) and Experian's Simmons Market Research Bureau (SMRB) measure the readership of large circulation consumer magazines and Scarborough Research measures the audiences of daily newspapers. Because readership data are usually unavailable for trade publications, cost efficiencies are usually calculated using the following cost-per-thousand formula: cost-per-thousand = advertising rate ($\times$ 1000)/circulation.

The circulations reported by trade publications are authenticated by several auditing services, the largest in the B2B marketing field being the Business Publications Audits of Circulation (BPA). The BPA primarily audits controlled circulation publications, which are funded entirely by advertising and provided at little or no cost to readers, but it also audits business publications that principally have paid circulations (e.g., *Marketing News*). The BPA is a nonprofit, self-regulating organization created by advertisers, publishers, and advertising agencies to assure that the circulation figures provided by business publications are accurate. It audits the circulations of more than 1900 business publications worldwide, including many publications for marketing professionals, such as *Adweek*, *BrandWeek*, and *B to B*, a 15-issue per year magazine for B2B marketing professionals. The audits also assure that direct mailings based on the trade publication's subscriber list also reach the intended targets.

Within trade publications, *Brandweek* and *B to B* are classified as horizontal trade publications. Horizontal trade publications run across industries, and usually have larger circulations than vertical publications. For example, *B to B* is read by marketing professionals in many industries, including the financial services, manufacturing, retail, travel, and publishing industries. Vertical publications are distributed with a single industry. As an example, *Progressive Railroading* is distributed to employees in the railroading industry, including professionals in management, marketing, signaling, purchasing, and even locomotive repair.

The BPA, when auditing trade publication circulations, analyzes and reports the industry and job classifications of subscribers, as well as their geographic locations, which allows

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

media planners to more efficiently compare the cost efficiencies of vertical and horizontal trade publications. The target group is referred to as the magazine's qualified circulation, and cost-per-thousands can be calculated using the qualified circulation. The circulation breakdowns also allow a B2B media planner to compare the cost efficiencies of reaching the qualified subscribers who work in the media (such as at broadcasting stations and newspapers) using a horizontal publication such as *Adweek*, as compared to vertical publications such as *Broadcasting & Cable* and *Editor & Publisher*.

Similar to the BPA, the Audit Bureau of Circulation (ABC), verifies the circulation statements of consumer magazines, newspapers, and some business publications, particularly those with large percentages of paid circulation. Like BPA, the ABC is a nonprofit, industry-sponsored organization. As an example, *Advertising Age*, which has more paid subscribers than nonpaid subscribers, is audited by the ABC. Some media planners tend to devalue nonpaid circulation when compared to paid circulation on the grounds that subscribers who pay to subscribe are probably more likely to read the periodical than those who receive it for free.

The ABC also conducts web audits. To assure that its estimates of web visits are accurate, the ABC requires online publishers to submit their log files to the ABC for analysis. The ABC analyzes the logs, removing visits generated internally, by robots and by duplication. The ABC publishes a "Consolidated Media Report" that combines circulation and web site traffic for trade publications, which can be used by media planners to calculate the relative cost efficiencies of the combined media. The ABC's consolidated report produced some controversy because the "total contacts" figures that it presents includes duplication – individuals who have read the magazine and have also visited the publication's web page, blogs and forums (Schwartz, 2006). Thus the combination does not produce an estimate of the vehicles' reach, but gives only its gross impressions. The BPA also audits visits to over 100 B2B websites. The BPA audit contains monthly data on unique browsers, page impressions, browser frequency, session duration, and page duration (BPA Worldwide, 2009).

The ABC has also started verifying readership studies conducted by trade publishers that use its auditing services. This is an attempt to make data for trade publications comparable to that for consumer magazines and newspapers, which often report readership rather than circulation numbers. The ABC works with the marketing research firm hired by the publisher to conduct its readership study, and assures that the sample selected is representative. The ABC also monitors the response rates and method for projecting the readership of the publication. Although relatively few trade publications measure readership when readership data are available, the cost efficiencies of vehicles can be assessed using the cost-per-thousand formula for comparing consumer magazine and newspapers: cost-per-thousand = rate × 1000/readership.

A few other companies, including Verified Audit Circulation (VAC), conduct comparable audits for niche trade and regional publications, direct to business marketing, and internet websites. VAC audits the circulations of publications such as *American Drycleaner*, *Idaho Business Review, Pit & Quarry*, and *Southern Loggin' Times*.

Media planners use the audit reports to compare one publication's circulation against another's. A publication with a higher qualified circulation may be seen as more valuable to an advertiser than a publication that has a higher unqualified circulation. Additional information gathered by an audit includes how a publication is addressed to its recipients, a geographical breakout of qualified circulation by state and region, and average qualified circulation over the previous 5 years. These data can be used by media planners to estimate whether a magazine's circulation has probably increased or decreased since the last audit.

In summary, a B2B advertiser's communication and marketing objectives can be obtained through the skilled use of tools available to the media planner. Auditing services play a valuable role in helping media planners select the proper vehicles to place advertisements. These services provide an equalized set of data which help media planners compare the merits of publications, digital media, and direct mail.

Bibliography

BPA Worldwide (2009) *Reading the BPA Worldwide Business Circulation Statement*. http://www.bpaww.com/Bpaww_com/Pages/MediaBuyer.aspx (retrieved on November 30 2009).

Schwartz, M. (2006) New ABC Audit for Business Publications Sparks Debate. *B2B*, December 11, p. 4.

Sorce, P. and Dewitz, A. (2007) The Case for Print Media Advertising in the Internet Age. A Research Monograph of the Printing Industry Center at RIT, pp. 1–37.

advertising effectiveness

Jacqueline J. Kacen

Introduction

In 2001, advertisers paid an average of $2 million for 30 s of television commercial time during Super Bowl XXXV (O'Connell, 2002). Of seven nationally known brands that bought advertising during the game – Budweiser, Cingular, Dentyne Ice, Doritos, MasterCard, Pepsi-Cola, and Visa – only two saw a gain in market share during the year (O'Connell, 2002). Does advertising work?

Advertising Effectiveness Defined

The goal of advertising is to persuade. *Advertising* is defined as any paid, mediated form of communication from an identifiable source, designed to persuade the receiver to take some action, now or in the future (Richards and Curran, 2002, p. 74). *Advertising effectiveness* is defined as an evaluation of the extent to which a specific advertisement or advertising campaign meets the objectives specified by the client (American Marketing Association, 2009). The effectiveness of advertising depends, of course, on what *specific* objective the advertiser has, and what *specific* action the advertiser is trying to persuade the receiver to take with respect to the advertisement.

At the most basic level, advertising effectiveness can be defined as some form of positive human response to a marketing communication. Human responses to advertising can be separated into two general categories: (i) effects on behavior, that is, sales, and (ii) cognitive and affective effects, that is, awareness, perceptions, and beliefs. Advertising, if successful at stimulating sales or influencing behavior, must first produce effects at the individual cognitive and affective level.

Difficulty Measuring Advertising Effectiveness

The effectiveness of advertising is difficult to measure. Establishing a direct link between exposure to mass media advertising (e.g., television commercials, billboards) and subsequent purchase is problematical (exceptions include *direct response advertising*). First, advertising is only one aspect of the "4Ps" (product, price, place, and promotion) that leads to sales. Consumers may buy a product for a variety of reasons. In fact, for most consumer-packaged goods, price and sales promotions have been shown to have a greater influence on purchase choice than exposure to advertising (Tellis, 2004). Some products achieve sales without advertising, due to word of mouth, news stories, and other nonpaid communications about the brand. Isolating the effect of advertising on sales while controlling for other factors is a difficult task.

Second, advertising has cumulative, lagged, and carryover effects. Repeated exposure to an ad may be necessary before interest in the product is generated. Exposure to an ad may occur long before the purchase takes place. In addition, the longer a brand is known, the more its ads are remembered (Kapferer, 2008). Brands that have been advertising for a long time (e.g., Coca-Cola, Budweiser) are familiar and may require little or no advertising in a given period to achieve sales results. Some of the long-term sales effects of current-period advertising can be attributed to consumers' earlier experiences and familiarity with the product (Givon and Horsky, 1990). Carryover effects in advertising mean that early exposure to an ad can influence later purchase decisions even when no advertising has occurred in that sales period (Tellis, 2004). Advertising makes positive brand evaluations and attitudes readily accessible in memory (Farquhar, 1989). As a result, advertised brands more readily come to mind even when no advertising is present.

Third, competitive activity impacts consumers' responsiveness to advertising. In some cases, advertisements for competitive brands can "prime" consumer interest in the whole category of brands, leading to greater advertising effectiveness. In other cases, competitive advertising can lead to advertising clutter, reduced share of voice, and decreased attention to advertising, reducing its effectiveness.

Findings on Advertising Effectiveness

On the basis of an extensive review of 50 years of research studies investigating the elasticity of advertising across various product categories, product life-cycle stages, countries, types of data,

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

and types of statistical models, the evidence suggests that a 10% increase in advertising leads to a 1% increase in sales (Tellis, 2004). This relatively small effect indicates that advertising is not an all-powerful force. As a point of comparison, the elasticity of price to sales is about 15% (Tellis, 2004). Advertising does influence behavior and generates sales, but it is a subtle force that must be carefully planned and executed to be effective.

How Advertising Works

The majority of advertising is brand-oriented, designed to persuade a target audience to respond favorably toward the advertised product– that is, to cause a change in beliefs, attitudes, and intentions in such a way as to motivate purchase. Traditionally, advertisers have viewed these mental processes in terms of a *hierarchy of effects* (see e.g., McGuire, 1978). The hierarchy-of-effects approach to advertising sees the individual as an information processor who proceeds through a linked chain of steps that begins with exposure to the ad, and is followed by attention to the brand, comprehension and understanding of the brand message, conviction about the brand (attitude change), and finally, purchase (McGuire, 1978). It is the job of advertising therefore, to promote the saliency (awareness) of the brand, to endow it with uniqueness and perceived differences from other brands in the marketplace, and to enhance the brand's reputation in the minds of consumers (Kapferer, 2008). Advertising achieves effectiveness by generating in consumers the mental processes that positively influence their beliefs and associations about the brand, enhance brand desirability, and encourage purchase.

Hierarchy of Effects

Awareness. Awareness of the brand is affected by the amount of brand advertising (*advertising weight*) and the placement of the advertising (where and when the ads appear), as well as the total amount of advertising vying for the individual's attention. High levels of advertising increase the likelihood that the brand message will break through the clutter of other advertising messages, and be seen by the target market. Research has demonstrated that companies should advertise heavily when the product is introduced but can reduce the amount of advertising later in the product's life cycle (Horsky and Simon, 1983). For familiar brands, expanding *reach* (number of people exposed to an advertisement) is generally more effective at stimulating sales than increasing *frequency* (the number of exposures a person has to an advertisement) (Ephron, 1995; Tellis, 2004).

While the level of advertising is important for creating brand awareness, the content of the advertising, the media through which the advertising is communicated, and when the advertising appears are critical to the hierarchy-of-effects process – that is, moving consumers from awareness to purchase. For example, research has shown that advertising is more effective when the brand name appears at the beginning of an ad rather than at the end (Baker and Honea, 2004). Attention to television commercials drops second by second; if the brand name is not mentioned until the end of the commercial, the memory association between the brand name and the message content is weakened, making the advertising less effective (Baker and Honea, 2004; Stewart and Koslow, 1989; Wells, 1997). Overall, advertising content, media channels, and scheduling generally have a greater impact on sales than the amount of advertising.

Persuasion. The main purpose of brand advertising is to persuade the consumer to purchase the brand. The attitude formation/change process differs depending on the consumer's level of involvement with the advertising (Krugman, 1966). Involvement plays a key role in determining how much effort is put into processing an ad message, and what kind of ad content (message arguments) is most persuasive. According to the *elaboration likelihood model* (see Petty, Cacioppo, and Schumann, 1983), under conditions of high involvement with the ad message, consumers engage in thoughtful processing of the information in the ad and generally require strong logical arguments in favor of the advertised brand in order to be persuaded to choose the brand (central route to attitude change). When there is low involvement, persuasion is aided by repetition and peripheral cues (like a celebrity endorser) that help increase the likability of the ad and generate positive perceptions of the brand

(peripheral route to attitude change). Under low involvement, attitude formation can occur after actual purchase (Krugman, 1968; Haley and Baldinger, 2000).

Attitudes, and changes in attitudes, are better predictors of sales than is mere recall of an advertisement (Wells, 1997). Research also indicates that attitudes changed via the central route to persuasion are more persistent over time, more resistant to change when attacked, and more predictive of purchase behavior compared to attitudes changed via the peripheral route (Petty, Cacioppo, and Schumann, 1983; Wells, 1997).

The optimal number of exposures needed to influence purchase choice depends upon a range of factors including whether the brand is an established or new product, brand dominance/market share, brand loyalty, length of purchase cycle, target market, message complexity, message uniqueness, new versus continuing campaign, clutter level, share of voice, scheduling (continuous or pulsed), and number and type of media (Ostrow, 1984). However, some research suggests that sales can occur after just one exposure to the ad message (Jones, 2002) and that, in general, the number of exposures needed is small – two to three exposures (Krugman, 1968; Tellis, 2004).

Creative content. The creative content of advertising is crucial to its effectiveness. Research in this area cannot provide a definitive set of guidelines as to what creative execution is most persuasive since the effectiveness of an advertisement depends upon many factors including brand familiarity, message complexity, and the novelty of the creative content. The attractiveness and credibility of the source, message elements such as type of appeal, strong or weak arguments, and other characteristics, such as length, size, color, layout, and sound, all impact the extent to which consumers are persuaded to purchase the brand (see Stewart and Koslow, 1989; Tellis, 2004). However, some general findings regarding creative content indicate that (i) advertising is more effective when the brand name appears at the beginning of an ad rather than at the end (Baker and Honea, 2004); (ii) rational appeals (strong arguments, expert sources) are more effective in new-product markets while emotion-based appeals are more effective in mature-product categories (Chandy *et al.*, 2001; MacInnis, Rao, and Weiss, 2002); (iii) celebrity endorsers are more effective under conditions of low involvement than high involvement (Petty, Cacioppo, and Schumann, 1983); and (iv) comparative advertising is most beneficial for new brands that compare themselves to established brands rather than the reverse (Grewal *et al.*, 1997).

Conclusion

Companies spend millions of dollars a year on advertising. According to the publication Media Matters (2007), the average person is exposed to about 600 ads per day. Political candidates want our votes, manufacturers and retailers want us to buy their brands, and nonprofit organizations want us to support their causes. Brands, ideas, and people constantly compete for our attention. Effective advertising breaks through this clutter and gets us to respond to the advertiser's message.

It is the job of advertising to persuade – to create clear, unique, and desirable associations that lead to preference for the brand and, ultimately, purchase. Effective advertising works through a hierarchy of effects that generates cognitive and affective changes in an individual's perceptions, beliefs, and attitudes. These mental processes enhance the advertised brand's value and stimulate sales. The process is not automatic. The ability of advertising to drive sales is tenuous and requires careful strategic planning and execution to be effective. However, effective advertising can build brand loyalty and lead to higher profitability, the generation of corporate goodwill, and reduced costs of marketing for the firm.

See also *attitudes; advertising message appeals; brand equity; consumer involvement; persuasion*

Bibliography

American Marketing Association (2009) Dictionary (available online 1 September 2009) at http://www.marketingpower.com/_layouts/ Dictionary.aspx.

Baker, W.E. and Honea, H. (2004) Do not wait to reveal the brand name: the effect of brand-name placement on television advertising effectiveness. *Journal of Advertising*, **33**, 77–85.

Chandy, R.K., Tellis, G.J., MacInnis, D., and Thaivanich, P. (2001) What to say when: advertising appeals in evolving markets. *Journal of Marketing Research*, **38**, 399–414.

Ephron, E. (1995) More weeks, less weight: the shelf space model of advertising. *Journal of Advertising Research*, **35**, 18–23.

Farquhar, P. (1989) Managing brand equity. *Marketing Research*, September, **1**, 24–34.

Givon, M. and Horsky, D. (1990) Untangling the effects of purchase reinforcement and advertising carryover. *Marketing Science*, **9**, 171–187.

Grewal, D., Kavanoor, S., Fern, E.F. *et al*., (1997) Comparative versus noncomparative advertising: a meta-analysis. *Journal of Marketing*, **61**, 1–15.

Haley, R.I. and Baldinger, A.L. (2000) The ARF copy research validity project. *Journal of Advertising Research*, **40**, 114–135.

Horsky, D. and Simon, L.S. (1983) Advertising and the diffusion of new products. *Marketing Science*, **2**, 1–17.

Jones, J.P. (2002) *The Ultimate Secrets of Advertising*, Sage, Thousand Oaks.

Kapferer, J.-N. (2008) *The New Strategic Brand Management*, Kogan Page, Philadelphia.

Krugman, H.E. (1966) The measurement of advertising involvement. *Public Opinion Quarterly*, **30**, 583–596.

Krugman, H.E. (1968) Processes underlying exposure to advertising. *American Psychologist*, **23**, 245–253.

MacInnis, D.J., Rao, A.G., and Weiss, A.M. (2002) Assessing when increased media weight helps sales of real world brands. *Journal of Marketing Research*, **39**, 391–407.

Media Matters (2007). Our rising ad dosage: it's not as oppressive as some think. *Media Matters*, February 15, 2007.

McGuire, W.J. (1978) An information-processing model of advertising effectiveness, in *Behavioral and Management Science in Marketing* (eds H.L. Davis and A.J. Silk), Wiley, New York, pp. 156–180.

O'Connell, V. (2002) Super bowl gets competition. *The Wall Street Journal*, **B1**, 3.

Ostrow, J.W. (1984) Setting frequency levels: an art or a science? *Journal of Advertising Research*, **24**, I9–I11.

Petty, R.E., Cacioppo, J.T., and Schumann, D. (1983) Central and peripheral routes to advertising effectiveness: the moderating role of involvement. *Journal of Consumer Research*, **10**, 135–146.

Reynolds, T.J., Olson, J.C., and Rochon, J.P. (1997) A strategic approach to measuring advertising effectiveness, in *Measuring Advertising Effectiveness* (ed. W.D. Wells), Erlbaum, Mahwah, pp. 337–355.

Richards, J.I. and Curran, C.M. (2002) Oracles on 'Advertising': searching for a definition. *Journal of Advertising*, **31**, 63–77.

Stewart, D.W. and Koslow, S. (1989) Executional factors and advertising effectiveness: a replication. *Journal of Advertising*, **18** (3), 21–32.

Tellis, G.J. (2004) *Effective Advertising: Understanding When, How, and Why Advertising Works*, Sage, Thousand Oaks.

Wells, W.D. (ed.) (1997) *Measuring Advertising Effectiveness*, Erlbaum, Mahwah.

brand equity

Jacqueline J. Kacen

Introduction

In 2009, the Coca-Cola brand was valued at $68.7 billion dollars and it was the world's most valuable brand (www.interbrand.com). The brand name alone, an intangible asset, accounted for 51% of the stock market value of the Coca-Cola Company (Interbrand, 2004). Coca-Cola is a brand with equity. Coca-Cola not only illustrates the value of a brand name but also demonstrates the importance of brands to a company's financial well-being.

The term *brand equity* came to public attention in the 1980s. It was first used by advertising practitioners to capture the idea that a brand had value to customers and because of this, the brand had financial value to the firm due to expected future sales (Barwise, 1993). In the late 1980s, quantifying the financial value of the brand became more important. Financial markets were recognizing the increasing gap between companies' book values and stock market valuations, as well as noticing the premium prices paid above stock market value during the mergers and acquisitions taking place. Large amounts of goodwill on balance sheets needed to be accounted for in an acceptable way (Feldwick, 1996; Murphy, 1990).

Definition of Brand Equity

Brand equity is the "value of the brand" (American Marketing Association, 2009). A brand is "a name, term, design, symbol, or any other feature that identifies one seller's good or service as distinct from those of other sellers" (American Marketing Association, 2009). In general, brand equity can be thought of as a set of brand assets and liabilities linked to a brand, its name and symbol, that add to or subtract from the value provided by a product or service to a firm and/or that firm's customers (Aaker, 1991). Value is assessed from both the customer and managerial (i.e., financial) perspective. From a customer perspective, brand equity is based on consumer attitudes about positive brand attributes and favorable consequences of brand use. From a managerial perspective, brand equity is the incremental cash flows which accrue to branded products over and above the cash flows which would result from the sale of unbranded products (Simon and Sullivan, 1993). From a financial markets perspective, brand equity is the sum of future profits, discounted in each period by a risk-adjusted interest rate, that result from associating that brand name with particular products or services (Simon and Sullivan, 1993).

Why Brands are Important

Customers like brands because brands provide assurances of quality, reliability, and consistency. Brands reduce perceived risk and allow people to shop quickly and confidently, saving time and energy on purchase decisions. Brands provide emotional rewards and experiential pleasure (Kapferer, 2008). Through brands, individuals can construct a self-concept and convey that concept to others.

Manufacturers benefit from brands as well. Brands cannot be copied by competitors, allowing manufacturers to command higher prices for preferred brands, obtain a larger market share, and spend less on promotional expenses such as advertising and price promotions (Shocker and Weitz, 1998). In addition, customer demand for the brand means manufacturers are better able to obtain distribution and cooperation from channel intermediaries. Customer preference for the brand can lead to a strong loyal customer franchise, more reliable demand forecasts, and expectations of cash flow from future sales. Last, a successful brand can be leveraged to increase sales and revenues for the firm through geographic market expansion and brand extensions.

Channel intermediaries, such as wholesalers and retailers, like brands because they allow for easier entry into markets, higher sales, and the opportunity to build loyalty at the store level.

How Brand Equity is Measured

The brand name and what it represents are a company's most important assets; a brand is the basis of competitive advantage and future revenue streams (Aaker, 1991). It is clear that brands help generate sales, and that sales

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

generate revenue and profits, and that long-term revenue is earned through repurchase and customer loyalty. It is also evident that the financial value of a brand is a consequence of the value customers place on the brand.

Measuring a brand's equity is a useful exercise to firms not only because it allows firms to assign a financial value to the brand for purposes of balance sheet accounting and other financial transactions but also knowing a brand's equity can help guide marketing strategy and tactical decisions, it can help managers assess the extendibility of the brand name to new products and new markets, it can help managers evaluate the effectiveness of their marketing decisions, and it can track a brand's health over time, and against competitive brands (Keller, 1993).

Many methods have been proposed to value a brand. These approaches are generally financially driven methods or research-based methods. Financially driven approaches include the cost-based value of the brand – that is, aggregate expenditures incurred developing the brand including research and development, marketing, advertising, and employee training costs (Aaker, 1991), the residual market value of the brand after other sources of value have been accounted for (Simon and Sullivan, 1993), the ratio of the brand's price to its competitor's price when both products are equally desirable to consumers (Crimmins, 1992), net present value of the price premium the brand commands over unbranded, generic or competitive brands (Aaker, 1991), acquisition value of the brand (Brasco, 1988), and net present value of the brand's future expected earnings (Brasco, 1988).

Market-level indicators of brand value – a loyal customer base, price premiums, and sustained market share leadership – are clearly the result of consumer-level effects. One may conceive brand equity as composed of consumer brand equity leading to market brand equity. The associations and beliefs the consumer has about the brand (brand image) and the strength of the consumer's attachment to the brand (brand loyalty) lead to positive market outcomes such as market share leadership, price premiums (brand strength) and ultimately, the financial value of the brand as a separate asset on a balance sheet (brand value).

Consumer-based research methods include measures of consumer perceptions, attitudes, awareness, knowledge, familiarity, preference, purchase intentions, satisfaction, loyalty, purchase share, and repurchase rates. The hierarchy of effects model is a useful framework for understanding customer-based brand equity (Agarwal and Rao, 1996). The hierarchy of effects model suggests that consumers pass through different stages, from exposure to the brand to intention to buy, before making a purchase. Consumer brand equity begins with brand awareness, and progresses to perceptions and associations relating to the brand, preferences for the brand, choice intentions, and ultimately, purchase.

Brand awareness can be assessed based on measures of unaided recall, recognition, or familiarity with the brand name. Without awareness, a brand cannot become part of the consumer's consideration set and no purchase can take place, so brand awareness is a critical first step to building brand equity. The associations and beliefs a consumer has about the brand comprise the brand image and represent a deeper level of brand knowledge. Brand image can be assessed through measures of perceptions of quality, specific beliefs about the brand on key attributes, and perceived differences with competitive brands. Highly valued brands are perceived as superior on some valued attribute. Central to building a valuable brand is the ability of the firm to develop and nurture clear, unique, and desirable brand associations. Measures of brand preference include the consumer's overall evaluation of the brand, consumer willingness to pay a premium for the brand, and loyalty or share of category purchases. Consumer purchase intention measures include likelihood of purchase or selection from a choice set. Consumer-based brand measures are often collected via survey or consumer panels and as a result, can be costly to obtain and are subject to the potential biases and errors inherent in survey research. However, consumer measures are necessary and critical antecedents to market-related measures of brand equity. A firm achieves higher market share or price premiums because of the perceived value of the brand to consumers.

Customer-based measures of equity lead to brand strength in the marketplace. Brand

strength can be thought of as consumer demand for the brand relative to competitive offerings. Brand strength is a behavioral outcome of the perceived value a brand holds for consumers. It is captured by measures of competitive superiority including market share leadership, market penetration, share of requirements, and price premiums. In essence, brand strength is evidenced in the increased volume and/or increased margins a firm enjoys compared to sales of other brands or unbranded products in the category.

Brand strength is a precursor to brand financial value – a monetized measure of brand equity such as the value of the brand on a balance sheet or the acquisition value of the brand.

Brand image, brand loyalty, brand strength, and brand financial value are not independent of each other. But at the same time, there is no empirical evidence to support an assumption of causal linkages between these four concepts. Brands can score high on consumer measures of awareness, preference, and purchase intention and yet perform poorly in a competitive marketplace. Establishing empirical links between consumer-based measures of brand value and financial-market based measures of brand value – especially future financial performance of the brand – remains a challenge for researchers in the area of brand equity.

Conclusion

Brand equity is an important concept. The concept of brand equity represents the combined effect of customer, firm, and competitive market decisions on the value of a brand. Firms that successfully leverage the value of their brands generate greater sales and profits. While brands are strategic assets and the source of competitive advantage for firms, brand equity is a measure of that success.

Bibliography

Aaker, D.A. (1991) *Managing Brand Equity: Capitalizing on the Value of a Brand Name*, Free Press, New York.

Agarwal, M.K. and Rao, V.R. (1996) An Empirical comparison of consumer-based measures of brand equity. *Marketing Letters*, **7** (3), 237–247.

American Marketing Association (2009) Dictionary, available online at http://www.marketingpower.com/_layouts/Dictionary.aspx (accessed on 1 September, 2009).

Barwise, P. (1993) Brand equity: snark or Boojum? *International Journal of Research in Marketing*, **10**, 93–104.

Brasco, T.C. (1988) How brand names are valued for acquisition, *Marketing Science Institute Conference, Austin, TX* and Marketing Science Institute, Cambridge, Report no. 88-104.

Crimmins, J.C. (1992) Better measurement and management of brand value. *Journal of Advertising Research*, **32**, 11–19.

Feldwick, P. (1996) What is brand equity anyway, and how do you measure it? *Journal of the Market Research Society*, **38** (2), 85–104.

Interbrand (2004) Brand Valuation: The Financial Value of Brands, April 27, available online at http: //brandcameo . org/ papers_review.asp?sp_id=357 (accessed on 30 September 2009).

Kapferer, J.-N. (2008) *The New Strategic Brand Management*, Kogan Page, Philadelphia.

Keller, K.L. (1993) Conceptualizing, measuring, and managing customer-based brand equity. *Journal of Marketing*, **57**, 1–22.

Murphy, J. (1990) Assessing the value of brands. *Long Range Planning*, **23** (3), 23–29.

Shocker, A. and Weitz, B. (1998), A perspective on brand equity principles and issues. *Marketing Science Institute Conference, Austin, TX*. Marketing Science Institute, Cambridge, Report no. 88-104.

Simon, C.J. and Sullivan, M.W. (1993) The measurement and determinants of brand equity: a financial approach. *Marketing Science*, **12** (Winter), 28–52.

advertising media selection

Kyle Krueger and Lawrence Soley

Advertising media selection is a major component of the media planning process performed by professionals in the advertising industry. It is the process of selecting the appropriate media (e.g., magazines) and vehicles (e.g., *Sports Illustrated*) to reach the target market (e.g., men, aged 25–54) of an advertising campaign. The process consists of making quantitative decisions, such as determining the most cost-efficient media and finding vehicles with a high composition of the target audience, as well as qualitative decisions, such as choosing media vehicles that provide an appropriate editorial environment for the product being advertised. Media planning is done to most efficiently achieve the media objectives of the campaign, such as exposing a percentage of the target audience to an advertisement or exposing the target audience to a certain number of advertisements.

Media objectives are frequently stated in terms of reach and frequency. Reach consists of the number of different prospects in the target market who will be exposed to the advertising campaign at least once. Reach is often stated as the percentage of the target market exposed to the campaign (e.g., 60%). Frequency refers to the average number of times each prospect will have the opportunity to be exposed to advertisements (e.g., 4.2 times) within a measured period of time. For branded products, it is believed that at least three exposures are necessary for a campaign to provide effective frequency. Generally, the first exposure generates some familiarity with the campaign message. The second exposure generates recognition; that is, that the message has been seen before. After the third exposure, the message begins generating recall (Krugman, 1972) (*see* ADVERTISING EFFECTIVENESS).

A summary measure of reach and frequency is called gross rating points (GRPs), which is the overall gross duplicated coverage of the target market that is reached at least one time, expressed as a percentage. GRPs is calculated as reach × frequency. Thus, a campaign generating 60% reach and a frequency of 4.2 would generate 252 GRPs.

GRPs is a simple way of estimating the number of exposures that an advertising campaign will generate, given the advertising campaign's budget (*see* THE ADVERTISING BUDGET). For example, prime time network television costs average $27 965 per rating point (SQAD, 2008). Thus, an advertising campaign with a budget of $20 million should be capable of buying at least 870 GRPs in prime time television. Media planners use this estimate as a starting point, attempting to find more efficient ways to reach the target market, as well as assessing what levels of reach and frequency are needed to achieve the communication goals of the campaigns.

Total impressions are another metric used to estimate the number of exposures generated by an advertising campaign. Unlike GRPs, which are expressed as a percentage, impressions are stated in terms of numbers of individuals (or homes). Total impressions can be calculated in two ways. The first method is to multiply the GRPs delivered by the population of the target group. A schedule that generates 252 GRPs against a target population of 1 000 000 prospects will produce 2 520 000 impressions. The second method is by adding the audience exposed (number) to each advertisement in the media schedule. Like GRPs, impressions signify the gross delivery of media schedule without regard to duplicate exposure (Surmanek, 1996).

To generate additional reach and frequency, media planners often use a media mix. A *media mix* refers to the use of two or more different media types in one advertising campaign. The most common reasons a media mix is employed are to reach prospects not reached by the first media type, to supply ancillary repeat exposure in a second medium after the optimal reach is achieved in the first medium, and to extend the efficacy of campaign creative through other media types conducive to the messaging. The basic assumption is synergistic when two media types are combined, the effect produced by the sum of parts is greater than the same two media used alone (Surmanek, 1996).

Another metric often used in the comparison of media vehicle or schedule efficiency is cost per thousand (CPM). CPM is the cost of delivering 1000 impressions against a defined audience group. For example, a full page ad

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

which costs $25 000 in a magazine with 500 000 subscribers has a CPM of $50, and is deemed more efficient than a similar sized ad that costs $15 000 in a magazine with 250 000 subscribers and has a CPM of $60.

Communication goals are statements about what the content of the advertising messages seeks to communicate. Communication goals are often stated in such terms as "prospects who recall the selling points of the product" or "become aware of where the product can be purchased."

Frequency is the foundation for media planning, as some advertising campaigns require higher levels of frequency to communicate effectively than others. If the needed frequency for a campaign is not generated, the campaign's communication objectives will not be achieved. Ostrow (1984) developed a model for estimating the needed frequency of an advertising campaign, which is sometimes known as the "*3M's*" of media planning. The 3M's are message factors, marketing factors, and media factors. Message factors consist of such things as the complexity of the message, the uniqueness of the message, and the length of the message. More complex messages, such as a message communicating multiple sales points, require higher levels of frequency than messages that have a simple message, such as generating brand name awareness. Although uniqueness is subjective, there is no question that unique messages stand out and require lower levels of frequency to get the message noticed and recalled than less unique messages. Similarly, campaigns using smaller print ads will require a higher level of frequency than smaller ads. Roper Starch studies have shown that full page ads generate greater attention than ads that are fractions of a page (Starch, 1966; Franke and Huhmann, 2004).

Marketing factors include the frequency of competitors' campaigns, the purchase cycle of the product, brand loyalty, and the characteristics of the products' target market. Generally speaking, if competitors have a heavy mass media presence, a higher level of frequency for a campaign is needed to break through this presence. Brands that have low brand loyalty need a higher level of frequency than brands that have high brand loyalty. Products that have a quick turnover, such as branded food products, need a higher level of repetition than products that are consumed less frequently, such as shoes. In addition, some target groups need higher levels of frequency than other target groups. Generally speaking, better-educated groups can process and retain advertising messages when exposed less frequently than less well-educated groups.

Media factors include the types of media used, the clutter in the media, position, and attentiveness to media. In terms of media, audio messages generally require a higher level of frequency than print or Internet messages, which allow prospects to reread messages if they fail to understand or remember a message component. Clutter refers to the number of messages surrounding the advertiser's message. Generally, the higher the clutter of the media, the higher the frequency needed for a campaign. Thus, higher levels of frequency are needed for campaigns using daytime television, which airs a greater number of commercials than prime-time television does. Moreover, campaigns that place ads in magazines that are highly cluttered or that bank ads need a higher level of repetition than campaigns that are not so cluttered. Banking is when magazines place advertisements together in specific parts of the magazine, separating them from the feature stories, as *National Geographic* does. Position refers to the location of the ad relative to others. Generally speaking, some positions, such as the covers and first five pages of a magazine, generate higher attention from readers than other positions. Additionally, ads positioned on the right-hand page are generally preferred over left-hand placements. Campaigns that are in these positions can use a lower level of frequency than those in other positions. Lastly, some media generate higher levels of attention than other media. Attentiveness to daytime television and reality shows is generally lower than attentiveness to prime time television and dramatic shows (Media Dynamics, 2009; Soley, 1984).

There are other ways to set frequency levels for campaigns. One is empirically, which means pretesting the campaign to determine the type and frequency of ads needed to achieve the communication goals of the campaign. There are commercial research services such

as Behaviorscan that will pretest messages for clients.

Much of the data for setting reach and frequency for media campaigns are available from syndicated research services, such as Simmons Market Research Bureau (SMRB) and Mediamark Research, Inc. (MRI). As an example, media planners can examine competitors' advertising activities using data supplied by such companies as TNS Media Intelligence and VMS or Internet-specific advertising activities using Nielsen NetRatings and comScore. Media planners can estimate levels of brand loyalty for brands with data supplied by SMRB. SMRB and Information Resources Inc. (IRI) also provide information on attentiveness to television dayparts and time spent reading magazines, allowing media planners to evaluate these media when setting advertising frequency levels.

In addition to providing data for making vehicle selection decisions, syndicated research firms such as MRI and SMRB can also be used for analyzing media consumption patterns among the target market, which is necessary for winnowing down available media types and the array of potential vehicles. This winnowing process is often done with a media quintile analysis. In a media quintile analysis, which segments people into a five-level spectrum of media consumption patterns, the indexes for a target group are compared against a base control group (which will have a 100 index). This comparison can provide media planners a starting point for media mix selection. For example, if the average adult uses television at a 100 index, and the heaviest users of TV among the target group (e.g., men, aged 25–54) use television at a 150 index, then the target group is 50% more likely to use television than the average adult. This medium would therefore be a potentially useful medium to include in the media mix. Indexes are also useful in winnowing down choices among media vehicles. For instance, the selection of radio formats can be assisted by the comparison of indexes. A format that has a 175 index against a target group is thought to be a better choice than a format that indexes at a lower number, outside the confines of cost and other planning parameters.

In summary, an advertiser's media objectives can be achieved through a proficient use of the tools available in the media planning process. Media planners combine quantitative date, such as syndicated research indices, and qualitative data, such as the education level of the target audience, in crafting a media that maximizes the efficiency of the media budget, and effectively targets the proper audience. A successful media plan combined with appropriate creative messaging leads to preference for an advertiser's brand and sales of an advertiser's products or services.

Bibliography

Franke, G.R. and Huhmann, B.A. (2004) Information content and consumer readership of print ads: a comparison of search and experience products. *Journal of the Academy of Marketing Science*, **32** (1), 20–31.

Krugman, H.E. (1972) Why three exposures may be enough. *Journal of Advertising Research*, **12** (6), 11–14.

Media Dynamics (2009) *TV Dimensions*, Media Dynamics, New York.

Ostrow, J.W. (1984) Setting frequency levels: an art or a science? *Journal of Advertising Research*, **24** (4), I9–I11.

Soley, L. (1984) Determinants of television program attentiveness. *Current Issues and Research in Advertising*, **7**, 141–148.

Starch, D. (1966) *Measuring Advertising Readership and Results*, McGraw-Hill, New York.

Surmanek, J. (1996) *Media Planning: A Practical Guide*, 3rd edn, NTC Business Books, Chicago.

SQAD. (2008) *Media Market Guide National Reports*, November 14.

emotional intelligence

Paula C. Peter

Introduction

Emotional Intelligence (EI) refers to the cognitive ability to correctly perceive, use, understand, and regulate emotions in the self and others (Mayer and Salovey, 1997). Although several definitions and measurements of EI exist, current literature in the field reinforces the existence of EI as a valid theoretical construct and supports its predictive validity in different domains.

Within the field of advertising and marketing communications, there is a well-established interest in the role that emotions play in advertising appeals. From the elaboration likelihood model (Petty and Cacioppo, 1986), to the use of sex, fear, and humor appeals in advertising, the past 30 years of marketing research and practice has highlighted the prominence of emotions in advertising appeals. However, there are relatively few studies that have examined factors that might moderate the persuasive effect of emotional appeals. EI might be considered as one of them.

EI, as the ability to correctly interpret and process emotional information, is expected to contribute to quality decision making as suggested by Goleman (1995) and supported by Damasio's work on neuroscience (Damasio, 1994). Emotionally intelligent individuals have a greater ability to recognize and understand emotions, which is necessary to differentiate the advertising appeal from the substance of the message. EI could benefit consumers in the interpretation of marketing communications, and could also benefit the marketers as they attempt to understand which emotional appeals are most effective in communicating messages about their product, service, or idea. The goal of this article is to provide an overview of EI in terms of its origin, theory, and measures in order to stimulate an area of research (EI) that has the potential to contribute to marketing communications and advertising.

A Historical Perspective

EI, already defined of the beginning was modeled in the early 1990s with work by Peter Salovey and John Mayer. Salovey and Mayer (1990) formed the basis for most of the academic research related to EI. They defined EI as *a cognitive ability that allows people to monitor, discriminate, and use one's own and other's emotional information in thinking and action*. They built their theory upon previous work on intelligence by Bar-On (1988, in Salovey and Mayer 1990), Gardner (1983, in Salovey and Mayer 1990) and Thorndike (1920, in Salovey and Mayer 1990). However, the construct did not receive much attention until 1995, when Daniel Goleman published his book "Emotional Intelligence: Why it can matter more than IQ". Although this book made the construct popular, it did not lead to a standardized definition of EI as originally formulated by Salovey and Mayer (1990); Mayer and Salovey (1997). In fact, the high level of interest in EI resulted in the creation of a variety of definitions of EI and what it might predict and also contributed to the formulation of different EI models.

Models of EI

The field of EI currently recognizes three basic models of EI: the ability model, the mixed model, and the trait model. However, the trait model is often considered a subset of the mixed model.

All models are founded on the basic premise of emotional intelligence but they differ in terms of definition and measurement of the construct. Emotional intelligence, as originally defined by Salovey and Mayer (1990), falls under the ability model of EI, where the focus is on emotional abilities measured by maximum performance on an ability test. In contrast, the mixed models and trait models focus on competencies, skills, and traits that are not measured through ability tests but through self-report measures instead (e.g., Schutte *et al.*, 1998).

The mixed model and trait model focus on a mix of constructs not necessarily related to emotions or intelligence. The mixed model focuses more on the prediction of EI, while the trait model focuses more on trait emotional intelligence within a personality framework. Work by Goleman (1998) on performance, and Bar-On

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

(2006) on health are examples of a mixed model approach toward EI. Goleman defines EI as a set of different competencies that drive success in the workplace. Bar-On defines emotional intelligence as an ability to cope with environmental demands and relate to ourselves and others. Schutte *et al.*, (1998) developed a self-report measure of EI that falls within the personality framework related to EI (trait model).

EI, as defined by the trait model and mixed model, has been criticized as not related to emotional competencies but rather to self-reported emotional traits instead. This criticism, when combined with a multitude of operalizations of EI, has contributed to a general skepticism toward emotional intelligence and its predictive validity. Self-reporting approaches to EI are considered to be appropriate measures of self-perceived EI, however, they are not appropriate to measure EI as cognitive ability.

In contrast, emotional intelligence as the cognitive ability to perceive, use, understand, and regulate emotions (Mayer and Salovey, 1997) has been conceptualized and measured using the ability model. Since the purpose of this article is to introduce emotional intelligence as an individual cognitive ability to correctly perceive, understand, and elaborate emotional information in advertising appeals, the focus will be on emotional intelligence as defined by the ability model (Mayer and Salovey, 1997).

The Ability Model of EI

The ability model represents the original model of EI and it was formally proposed by Mayer, DiPaolo, and Salovey (1990) after the conceptualization of EI by Salovey and Mayer (1990). The model was subsequently refined by Mayer and Salovey, 1997. Mayer and Salovey (1997) proposed EI as a cognitive ability that can vary from individual to individual. This cognitive ability is further subdivided into four specific branches, which are organized hierarchically:

1. Perception of emotions – the ability to recognize and identify emotions in the self and others.
2. Use of emotions – the ability to use emotions to facilitate thoughts and solve problems.
3. Understanding of emotions – the ability to understand and interpret blends of emotions.
4. Regulation of emotion – the ability to regulate emotions in self and others.

Measures of EI. Theoretical and empirical research on the four branches of cognitive abilities resulted in the development of the most widely recognized measure of EI: The Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT). The MSCEIT is a proprietary cognitive test of EI designed for people aged 17 and older and modeled on ability-based IQ tests. It consists of 141 questions and provides respondents with 3 different scoring subsets: an overall EI score; "area" EI scores; and "branch" EI scores. The two area EI scores are defined as the experiencing area of EI (formed by the perception and use of emotions), and the reasoning area of EI (formed by the understanding and managing of emotions). The branch EI scores measure the perceiving, using, understanding, and managing cognitive abilities previously outlined.

The MSCEIT interpretation of EI at these three major levels has shown good reliability, face validity, and discriminant validity with different measures associated with EI (Mayer, Salovey, and Caruso, 2002). The MSCEIT assesses emotional abilities with answers deemed correct by consensus score or by expert score.

Using the MSCEIT as a foundation, Kidwell, Hardesty, and Childers (2008) developed the Consumer Emotional Intelligence Scale (CEIS). The CEIS is a nonproprietary context-based measure of EI that relies on the same EI cognitive branches as originally defined by the MSCEIT. The scale is shorter (18 items vs 141) and offers an alternative to MSCEIT of assessing emotional abilities in the consumer domain. The CEIS has been related to higher quality decision related to food and general product choices (Kidwell, Hardesty, and Childers, 2008) but is not meant to be a substitute for more general measures of EI (i.e., MSCEIT). Owing to the relative newness of the CEIS, further research is required to validate the reliability, face, and discriminant validity of CEIS in the consumer domain.

What does EI predict? Mayer, Roberts, and Barsade (2008) have recently offered a comprehensive review of studies that examined the predictive validity of EI as measured by the MSCEIT. On the basis of their analysis, they suggest that EI is positively related to better social relations for children and adults, social acceptance, family and intimate relationships, academic achievement, social relations during work performance and in negotiations, and psychological well-being. They also propose EI to be positively related to career interests, attitude toward money, money gained in negotiations, emotional eavesdropping, and emotional forecasting (Mayer, Roberts, and Barsade, 2008).

The authors suggest the importance of maintaining proper controls for age, gender, cognitive intelligence (IQ), and personality variables while making predictions of EI in these domains. Otherwise, empirical evidence may reveal contradictory findings.

Can we learn and teach EI? EI competencies build on emotional knowledge, and emotional knowledge seems to be relatively easy to obtain and promote (Mayer, Salovey, and Caruso, 2004). However, very few studies provide empirical evidence of this phenomenon. This is probably due to the novelty of the construct, as well as a lack of clear understanding of how EI training programs should be developed. Moreover, the tendency is to use short-term training programs to test for the possibility of learning and/or teaching EI when a longer term perspective would be more appropriate. Mayer, Salovey, and Caruso (2004) provide an overview of several longitudinal studies and conclude that a developmental trend in EI exists and that further research is warranted in this area.

Criticism and future of EI. A great deal of criticism about EI has been related to its elusive and broad conceptualization, as well as its association with self-report measures. However, it is important to note that these specific criticisms do not apply to the ability model of EI, as this model is deeply rooted in psychology and has a well-established theoretical and empirical framework. Nevertheless, other legitimate criticisms of the ability model do exist.

First, the MSCEIT has been questioned as genuine measure of intelligence since it is not measured by objectively correct responses. As mentioned above, the MSCEIT offers a scoring based on answers deemed correct by consensus score (the norm) or by expert score. The option of consensus means that the correct response is the one the majority of the sample (N = 5000, Mayer, Salovey, and Caruso, 2002) has endorsed. Therefore, the creation of questions that only high emotional intelligent individuals can solve becomes impossible by the consensus criterion. The expert scoring might be suggested as an alternative to avoid criticism related to the consensus criterion. In the expert scoring, the correct answer is deemed correct by a pool of researchers who have expertise in emotions.

Second, research to date on EI lacks consistency in terms of the predictive validity of EI. A more systematic approach to EI research must be developed in order to determine the predictive validity of EI. Further, the relationship between EI and its areas, branches, and criterion variables should be hypothesized on the basis of a theoretical framework that calls for the salience of the cognitive ability to perceive, use, understand, and regulate emotions.

The validity of EI relies on its nomological framework, which includes a well-established theoretical (ability model of EI) and empirical (MSCEIT) framework. The future of EI, as a field of research, depends on a systematic approach toward the understanding of the linkages among the theoretical and empirical framework of EI.

Applications of Emotional Intelligence in Advertising

The majority of research and practice related to advertising and marketing communications suggests that emotional appeals are effective in attracting consumers' attention and may even build brand connections. Thus, marketers are interested in the identification of the appropriate evoked emotions to be linked to the brand, product, service, or idea. If marketers tap into the appropriate emotional connections, they stand a much better chance of developing a long-lasting relationship between their brand and its users. Conversely, a failure to create the appropriate

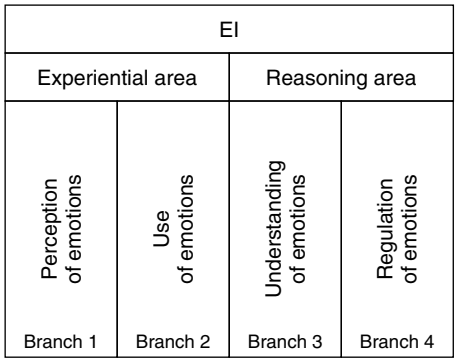


Figure 1 An illustration of the Ability Model of EI as measured by the MSCEIT.

emotional connections could lead to less of an impact and potentially drive consumers away, diminishing the brand value of their product, service, or idea.

Marketers could also take advantage of studies relating to how consumers use emotional information to actually make brand, product, or service decisions. EI, as defined by the ability model, represents a reliable tool for measuring individual differences in consumers' ability to use and process emotional information.

EI must be viewed as a multidimensional construct (Figure 1). For example, the perception of emotions (Figure 1, Branch 1) might be important in order to understand which features of the product emotionally appeal to the consumer, while the understanding of emotions (Figure 1, Branch 3) could be useful in understanding which emotions are most meaningful to consumers when they establish a connection with a specific product or brand.

These are only a few possible suggestions related to the application of EI in advertising and marketing communications. EI has the potential to contribute to a well-established literature that has considered both the role of emotions and logic in decision making (e.g., Damasio, 1994). Research on the role of EI would likely benefit the marketer and the consumer, as well as increasing the effectiveness of advertising and marketing communications.

Bibliography

Bar-On, R. (2006) The bar-on model of emotional-social intelligence (ESI). *Psicothema*, **18**, 13–25.

Damasio, A. (1994) *Descartes' Error: ~Emotion, Reason, and the Human Brain*, Grosset/Putnam, New York.

Goleman, D. (1995) *Emotional Intelligence*, Bantam Books, New York.

Goleman, D. (1998) *Working with Emotional Intelligence*, Bantam Books, New York.

Kidwell, B., Hardesty, D., and Childers, T. (2008) Emotional calibration effects on consumer choice. *Journal of Consumer Research*, **35**, 611–621.

Mayer, J.D., DiPaolo, M.T., and Salovey, P. (1990) Perceiving affective content in ambiguous visual stimuli: a component of emotional intelligence. *Journal of Personality Assessment*, **54**, 772–781.

Mayer, J., Roberts, R., and Barsade, S.G. (2008) Human abilities: emotional intelligence. *Annual Review of Psychology*, **59**, 507–536.

Mayer, J.D. and Salovey, P. (1997) What is emotional intelligence? in *Emotional Development and Emotional Intelligence: Educational Implications* (eds P. Salovey and D. Sluyter), Basicbooks, Inc, New York, pp. 3–34.

Mayer, J.D., Salovey, P., and Caruso, D.R. (2002) *Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT)*, Multi-Health Systems, Inc., Toronto, Ontario.

Mayer, J.D., Salovey, P., and Caruso, D.R. (2004) Emotional intelligence: theory, findings, and implications. *Psychological Inquiry*, **15**, 197–215.

Petty, R.E. and Cacioppo, J.T. (1986) *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*, Springer-Verlag, New York.

Salovey, P. and Mayer, J.D. (1990) Emotional intelligence. *Imagination, Cognition and Personality*, **9** (3), 185–211.

Schutte, N.S., Malouff, J.M., Hall, L.E. *et al.*, (1998) Development and validation of a measure of emotional intelligence. *Personality and Individual Differences*, **25** (2), 167–177.

brand extensions and flanker brands

Youngseon Kim and Tina M. Lowrey

Brand Extensions

A brand extension involves a company using an existing brand name to introduce a new product (e.g., Keller, 2008). That is, it is one of many new-product launch strategies. A brand extension stretches a well-established brand name for a new-product offering into either a totally different product category or in the same product category for a new market segment. The existing brand is called the parent brand or the core brand because it gives life to a brand extension (Keller, 2008). This implies that brand extensions fall into two general categories: category extension and line extension (Farquhar, 1989). Category extension occurs when a company uses the parent brand to launch a new product in a different product category from the one that it currently serves (e.g., Fendi watches, Jeep strollers, and Honda lawn mowers). Line extension occurs when a company applies the parent brand to a new product that targets a different market segment within a product category that the company currently serves (e.g., Ralph Lauren purple label and Head & Shoulders® dry scalp shampoo).

Brand extension has been a popular strategy for continued growth for major companies that already have a strong brand in the markets they serve. Recently, 82% of new product launches were identified as either type of brand extension (Simms, 2005). This trend may be because brand extensions are low-risk models in which a well-built brand name and its image can be transferred to a new product, compared to the high risk of launching a new product under a totally different brand name. Thus, brand extensions buffer the costs and risk associated with launching a new product. A familiar and trusted brand name signals credibility and quality to consumers and elevates the likelihood that they will try the extension product (Braig and Tybout, 2005). In general, brand extensions can help build/maintain a company's brand equity. They leverage brand assets and create synergy by generating more chances for brand exposure and associations in different contexts. In addition, they increase profitability by serving the customer with new products in more than one product category (Aaker and Joachimsthaler, 2000). Thus, well-planned and well-implemented extensions not only bring out a number of advantages to marketers in a new product introduction, but also benefit the parent brand.

A vertical extension is a different type of brand extension which does not simply fit into either category of brand extensions mentioned above. It includes sub-branding and super-branding targeted at different market segments in the same category. In this sense, vertical extensions can be said to belong to line extensions except that they have an additional brand name endorsed by the parent brand (e.g., Courtyard® by Marriott). Marketers introduce low-priced versions of their established brand name products for more value-conscious segments (sub-branding) whereas they attempt to ladder up their brand to up-market through brand extensions in order to serve a more premium market segment (super-branding) (Keller, 2002). The rationale behind vertical extensions is that the parent brand's equity can be transferred in either direction in order to appeal to consumers because the parent brand has a role as endorser for the new offering (Keller, 2008).

Sub-branding (e.g., Fairfield Inn® by Marriot, Marc by Marc Jacobs) can be used as a means to distinguish lower-priced entries. Sub-branding can reduce the risk of potential failure of the extension. However, downward extension of prestige brands can face resistance from current owners because they desire brand exclusivity (Kirmani, Sood, and Bridges, 1999). Super-branding can be used to indicate a noticeable, although presumably not dramatic, quality improvement (e.g., GE Profile™ or Levi's® Vintage). The upward extension can upgrade brand image because a more premium version of a brand tends to develop positive associations with the brand (Keller, 2008).

Considering the high failure rate of new products (e.g., only one or two out of ten new products are successful and the vast majority of them are withdrawn from the market within a year; Braig and Tybout, 2005; Keller, 2008) brand extensions definitely make new-product

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

acceptance easier if the parent brand has been well known and well liked in the market that it has served (Keller, 2008). Consumers can make inferences and expectations about the likely quality and performance of a new product, based on what they already know about the parent brand (the degree to which they feel the knowledge is relevant to the new product) (Keller, 2008). This benefit enables a reduction of risk perceived by customers, distributors, and retailers. It can lead to a decrease in the cost of gaining distribution and trial, improve the efficiency of promotional expenditures, and so on. In addition to facilitating acceptance of extensions, they can also provide feedback benefits to the parent brand and the company as a whole. They can enhance the parent brand image by improving the strength and favorability of brand associations. The enhancement of the parent brand also leads to improvement of perceptions of company credibility, expertise, and trustworthiness (Aaker and Keller, 1990). In a long-term perspective, extensions may help fine-tune the company's core benefit proposition and business definition, which is a key factor in gaining new customers in the company's brand franchise and increasing market coverage.

Despite the many number of benefits of brand extensions mentioned so far, they are not always easy to execute in the competitive business environment. Thus, marketers should be able to address questions ranging from whether to extend the brand to when, where, and how to extend it before making any strategic decision about launching a brand extension. Poorly executed extensions can dilute brand associations, weaken brand power, and reduce the clarity of the offerings. When a brand is tied closely to a certain product class and represents it, the brand's potential to stretch is limited. For example, brands such as Kleenex® and Clorox® Bleach may not extend too far beyond their basic product areas because they are so attached to a specific product and its attributes (Aaker and Joachimsthaler, 2000). In contrast, brands with credibility and intangible associations are more likely to extend to new categories because those intangibles make sense to consumers in a wide variety of contexts (Park, Milberg, and Lawson, 1991; Schultz and Schultz, 2004). Extreme examples of this can be found in Virgin brand. The brand has been stretched from record sales to colas to hotels and airlines. How were those extensions possible? It does not seem to represent expertise in one specific field; instead it may be a way of managing, a way of serving customers, and a view of the world that people seem to like (Schultz and Schultz, 2004). In contrast, Park, Milberg, and Lawson (1991) showed that function-based brands have more difficulty with brand extension than prestige-based brands (e.g., Timex vs Rolex).

What are the basic guidelines about brand extensions? Successful brand extensions can be made when the parent brand is viewed to have favorable associations and there is a perception of fit between the parent brand and its extension (Keller, 2002). The fit can be found in many different parts of a parent brand and its extension, ranging from not only product-related attributes and benefits, but also non-product-related attributes and benefits such as common usage situations or user types (Aaker and Joachimsthaler, 2000). Wherever the base of fit judgment is located, consumers should be comfortable with the extension and sense a fit. The brand extension should offer an added value for consumers, thereby helping consumers understand why the new offering from the parent brand should be preferred to other brands (Aaker and Joachimsthaler, 2000). In addition, the new association between the parent brand and its extension should enhance brand equity. Marketing research should be conducted to explore three of these criteria for making a decision of whether the brand can be extended successfully (Keller, 2008).

Academic research and industry experience have discovered a number of principles leading to successful brand extensions. Keller (2008) recommends that marketers go over their brand extension strategies thoroughly by following the steps below, and apply their managerial judgment and consumer research findings in order to accomplish appropriate and successful brand extensions.

- Define actual and desired consumer knowledge about the brand (e.g., create mental map and identify key sources of equity).
- Identify possible extension candidates on the basis of parent brand associations and overall

similarity or fit of extension to the parent brand.
- Evaluate the potential of the extension candidate to create equity.
- Design marketing programs to launch extensions.
- Evaluate extension success and effects on parent brand equity. (Keller, 2008, p. 524).

Flanker Brands

When a brand is attacked by a competitor with a value offer or unique position, any company takes such an attack seriously and responds. However, any careless response can jeopardize its image and brand equity (Aaker, 2004). A solution can be found through having a flanker brand. A flanker brand (also called a fighting brand) is a new brand launched in the market by a company with an established brand in the same product category. It is designed to fight a competitor, shielding the flagship brand from the fray (Giddens, 2002). The name "flanker brand" comes from a war metaphor. A flanker brand protects the flagship brand from a competitor that is not competing directly with attributes and benefits that the flagship brand has nurtured (Aaker, 2004). Ideally, a flanker brand should compete in the same category as the flagship brand without cannibalizing the flagship brand's market share through targeting a different group of consumers. The objective of a flanker brand is to debilitate the competitor brand where it is positioned without compelling the flagship brand to divert its focus (Aaker, 2004). Broadly this strategy is called fighter branding or multibranding in the sense that it can help a company to occupy a larger total market share than one product could garner alone.

The brand portfolio of a company with flanker brands includes the following types of offerings: (i) a flagship (or premium) brand that offers high quality at a higher price; and (ii) one or more "value" brands offering a slightly lower quality or a different set of benefits for a lower price. One of the largest companies to use this strategy effectively is Proctor & Gamble (P&G™). Tide® is an extremely well-established laundry detergent whereas Cheer® is a slightly lower quality detergent available at a value price, developed to appeal to consumers who wanted a lower cost detergent. The sales of Tide® sales dropped slightly when the new brand, Cheer®, was introduced, but the combined sales of Cheer® and Tide® were greater than the former sales of Tide® alone. Thus, the adoption of a flanker branding strategy by P&G™ let the company achieve a higher market share (Giddens, 2002).

Typically, the purpose of flanker brands is to create stronger points of parity with competitors' brands so that more important (and more profitable) flagship brands can retain their desired positioning (Keller, 2008). Many companies are launching discount brands as flankers to better compete with store brands and private labels, and guard their higher end brands. Thus, it is true that developing a flanker brand offers a number of advantages to the company adopting the strategy. For example, a flanker brand helps gain more store shelf space for the company which leads to retailers' increased dependence on the company's products (Giddens, 2002). At the same time, a flanker brand helps mitigate the pressure of losing sales caused by cheaper store brands or private label brands available in almost every retail store chain, and grab "brand switchers" for a value offering by providing several different brands. Above all, the company and the flagship brand can be protected even if the flanker brand fails in the market because consumers will not readily associate it with the existing flagship brand.

However, flanker branding is not completely without risk. The risk is similar to that of launching a new brand. Introducing a new brand into the market is always expensive. Brand introductions need an ample amount of marketing communication expenditures as well as a great deal of consumer research. Thus, in some cases, companies try to reposition existing brands in their portfolios to play the flanker role instead of introducing a new brand. For example, P&G™ repositioned its one-time top-tier Luvs diaper brand to serve as a flanker to insulate the premium-positioned Pampers brand from private labels and store brands (Keller, 2008).

When marketers design flanker brands, they must be very careful. Fighter brands must not be so attractive that they take large amounts of sales away from their higher priced comparison

brands. At the same time, flanker brands should not be designed so cheaply that they negatively affect the other brands in the company's brand portfolio (Keller, 2008). Like brand extensions, flanker branding may not be for every player in the market, even if it may enable a company to protect its flagship brands from the attack of lower priced value offerings by the competitor.

Conclusion

A brand portfolio strategy guides how a company will manage different brands and branding elements to lead to long-term profitable growth (Calkins, 2005). Brand extensions and flanker brands can be said to be key players in building and managing a strong brand portfolio, viewed from the objectives and roles of each. A growth alternative is found when utilizing a strong brand through extensions (Aaker, 2004). Brand extensions involve leveraging tangible and intangible assets of a parent brand (or a core brand) for introducing a new product. Whether a company has a strong core brand is the first and most important base for stretching a brand into a new market. Unless a company has a strong and well-built brand to extend, it may be fruitless to invest in brand extensions. Thus, building a strong core brand is the first priority before pursuing any type of extensions. Once the strong core brand is built, extension opportunities are easy to find, and a brand extension from the strong core brand will add value through its favorable associations and reach current/prospective customers (Aaker, 2004; Calkins, 2005). A company also makes a strategic decision to move down market in order to access volume markets. This movement may result in an unsuccessful launch of a new product, as well as damage the company's brand image. In terms of maintaining a well-conceived brand portfolio, the company can separate a value offering from its core brand through sub-branding. However, the best option may be to reposition an existing brand in the company's brand portfolio, or to create a new one (Aaker, 2004). This movement is in line with flanker branding. With flanker brands, the company is more likely to protect its core market position without diluting the core brand image and garner more sales.

The role and business objectives of brand extensions are very different from those of flanker brands. Each of them is rooted in a different branding strategy in a company's brand architecture and portfolio. However, they are equally important in terms of protecting the brand equity that the company has achieved by bringing out greater customer share on the condition that they are well designed and implemented. Finally, strategic utilization of brand extensions and flanker brands in a company's brand portfolio will help achieve a company's business goal of long-term profitable growth by meeting the following marketing objectives based upon special roles of brands in the brand portfolio of a company: (i) to attract a particular market segment that is not currently being served by other brands of the company (Keller, 2008); (ii) to keep market position and protect flagship brands (Giddens, 2002); (iii) to broaden product offerings in order to appeal to consumers seeking variety who may otherwise have switched to a competitor's brand (Keller, 2008); and (iv) to achieve economies of scale in marketing communications, sales, merchandising, and physical distribution (Keller, 2008).

See also *brand strategy; brand value; launch strategies; marketing strategy implementation; perception of brand equity*

Bibliography

Aaker, D.A. (2004) *Brand Portfolio Strategy: Creating Relevance, Differentiation, Energy, Leverage, and Clarity*, Free Press, New York.

Aaker, D.A. and Joachimsthaler, E. (2000) *Brand Leadership*, Free Press, New York.

Aaker, D.A. and Keller, K.L. (1990) Consumer evaluations of brand extensions. *Journal of Marketing*, **54**, 27–41.

Braig, B.M. and Tybout, A.M. (2005) Brand extensions, in *Kellogg on Branding* (eds A.M. Tybout and T. Calkins), John Wiley & Sons, Inc., Hoboken, pp. 91–103.

Calkins, T. (2005) Brand portfolio strategy, in *Kellogg on Branding* (eds A.M. Tybout and T. Calkins), John Wiley & Sons, Inc., Hoboken, pp. 104–125.

Farquhar, P. (1989) Managing brand equity. *Marketing Research*, **1**, 24–33.

Giddens, N. (2002) Building your brand with flanker brands, Ag Decision Maker, June, http://test.agmrc.org/ agmrc/ business/ operatingbusiness/ bybrand flankerbrands.htm

Keller, K.L. (2002) *Branding and Brand Equity*, Marketing Science Institute, Cambridge.

Keller, K.L. (2008) *Strategic Brand Management: Building, Measuring, and Managing Brand Equity*, 3rd edn, Pearson Prentice Hall, Upper Saddle River.

Kirmani, A., Sood, S., and Bridges, S. (1999) The Ownership Effect in Responses to Brand Line Stretches. *Journal of Marketing*, **63**, 88–101.

Park, C.W., Milberg, S.M., and Lawson, R. (1991) Evaluation of brand extensions: the role of product feature similarity and brand concept consistency. *Journal of Consumer Research*, **18**, 185–193.

Schultz, D.E. and Schultz, H.F. (2004) *Brand Babble: Sense and Nonsense about Branding*, South-Western, Mason.

Simms, J. (2005) Where are all the new idea? *Marketing (UK)*, Issue (December 28), 34–36.

advertising and the integrated marketing communications (IMC) process

Youngseon Kim and Tina M. Lowrey

Integrated marketing communications (IMC) involves strategic synchronization of various forms of persuasive communication programs with customers and prospects over time (e.g., Schultz, 1993). The goal of IMC is to convey a company's marketing messages in a consistent and coherent manner through online and off-line communication channels. Thus, IMC deals with all sources of brand or company contacts that a customer or prospect has about the product or service as potential delivery channels for conveying messages. Advertisers utilize a myriad of tools such as advertising, public relations, direct marketing, interactive/Internet marketing, sales promotion, and personal selling in order to best communicate with their target audience. Advertising in traditional media played a key role in providing and managing the initial attempts at integrating different marketing communication tools in the 1990s (Percy, 2008). This key role seemed legitimate, considering the size of the budget allocated to advertising. However, since then, there have been shifts in the advertising and media industries from conventional advertising methods to multiple forms of communications, from mass media to more specialized (niche) media, which are centered on specific target audiences, and from advertiser-driven message transfer to consumer-driven message contagion, and so on. Thus, the strategic importance of IMC has become more widely recognized. In addition, the role or value of traditional mass-media advertising today is often questioned.

However, IMC planning and implementation does not necessarily shrink the role or value of advertising in general (Percy, 2008). Advertising was primarily placed in traditional mass-media such as television, radio, newspaper, magazines, outdoor, and direct mail. Nowadays, advertising messages are increasingly delivered through co-op marketing programs (i.e., cooperative marketing communication programs between a manufacturer and its retailers, or co-op advertising/joint promotions between two brands. For example, when a company launches a campaign for a new brand, it tries hard to take advantage of the venues it can leverage in order to keep the new campaign running to communicate integrated and consistent messages at any possible customer contact point) and on the Internet. We can also find some advertising campaigns that contain a well-pronounced brand-building message as well as a promotional offer for a new product, all built to create an immediate excitement to buy (Percy, 2008). These are a few examples of how advertising effectiveness can be enhanced when paired with other promotion tools. They also show how advertising messages can be delivered more effectively and more efficiently throughout a fuller line of media channels, depending on the target audience and the marketing communication objectives.

The primary goal of advertising is to build brand awareness and brand attitude, which are fundamental factors for building a brand (e.g., Percy, 2008). If this is the company's major communication objective, advertising may lead the whole marketing communication program. Building a brand takes time. Thus, this is a long-term project. All other marketing communications should help build brand awareness and develop a positive feeling for the brand. That is the reason that IMC process and implementation are so critical in building strong brands. Brand messages should be developed to fit with the communication objective and should be conveyed in a consistent "look and feel" across media platforms at every consumer contact point over time, which ensures that everything included in an IMC campaign will send the same message (Percy, 2008). The key themes and visuals contained in advertisements should be embedded into other forms of marketing communication channels: public relations, direct mail, Internet presence, in-store communication, and so on. One of the most important and difficult tasks for IMC is ensuring consistency in executions so that the target audience is able to immediately identify any execution within a campaign and connect it to the brand in the end. In this cluttered and fragmented media environment, advertising in traditional media may have lost its power. However, it can still shine as a part of a well-defined IMC campaign.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

See also *brand strategy; integrated marketing communication strategy; integrated marketing communication; marketing strategy implementation; perception of brand equity; subliminal advertising*

Bibliography

Percy, L. (2008) *Strategic Integrated Marketing Communication Theory and Practice*, Butterworth-Heinemann, Burlington.

Schultz, D.E. (1993) Integrated marketing communications: maybe definition is in the point of view. *Marketing News*, **27**, 17.

luxury brands versus other brand categories

Kapareliotis Ilias

Brands are symbolic entities and powerful elements influencing cultures, societies, and generations. They play a pivotal role in everyday life from the different decisions to the choices people make every day. The branding concept appeared hundreds of years ago. Brands at that time meant differentiation from one cattle farmer to the other. The current brand definition has evolved from marks, names, logos, and product shapes. The definition evolved as the marketing strategies related to the brand were more focused to them. A brand symbolizes the guarantee and credibility that the consumer asks when purchasing. The brand is the final promise that a consumer will take after the final purchasing decision. The brand can thus be defined as follows: *A brand is a name, term, sign, symbol, design, or a combination of these that is intended to identify the product of a seller and to differentiate it from those of competitors*.

Branding has become very prominent over the last thirty years as a core marketing aspect and an important part of business strategy. As a modern concept, branding can be applied to anything from products to services, to companies, countries, and even individuals. Every brand has a unique characteristic of identity, promise, value, and differentiation. Additionally, there are brand elements that build the relationship between the consumer and the brand. The benefit is mutual for the consumer and the company as they both acknowledge the identity of each other. Consumers are delivered a certain promise not only of quality, but also of a unique experience. Thus the brand becomes not only a marketing tool, but also an intangible asset that guarantees profitability and future earnings.

Luxury goods have been developed over the last twenty years. The overall development of luxury brands is due to different factors. Some of them are due to the increase of income level in western societies, the need for differentiation, and the saturation of branded goods of the product category. Luxury brands give emphasis to the art of product design, innovation, and aesthetics. Branding for luxury products is based on design and creativity. Luxury brands are striving for innovation, differentiation, and uniqueness. Luxury brands set the fashion trend for every category they belong to. Luxury brands try to combine the past and the future. The luxury brand logos combine luxury images and prestige. Consumers expect luxury brands to be innovative in all aspects and be trendsetters. Luxury brands are more customized to consumers' needs than regular brands. Luxury brands try to understand consumer's psychology, changing habits, and needs, and of course, have to be in the forefront of everyday life. The producers of luxury brands need to make an analysis of current trends and social or other changes influencing consumers and to produce products that will attract as many consumers as possible.

Luxury brands have been crafted through consistent and sophisticated strategies in the marketing field. They have a high acceptance from the global world and in special moments, cases, or circumstances consumers think that luxury brands might be the remedy for their conscious or unconscious problems. Luxury brands have some main core characteristics that differentiate them from the other brand categories. The differentiating characteristics of luxury brands in comparison with other brand categories are *strong brand image, differentiation, innovation, exclusivity, high quality, and high pricing*.

A Historical Perspective

Branding was part of ancient trade traced back to Greek and Roman empires, long before the modern industrial revolution. Markets, during that period, were creating awareness through signs and product descriptions on stone footpaths. In the sixteenth century, a new method evolved when whisky distillers started to ship their products in wooden barrels with the name of the producer etched on the barrels. Thus the product and the maker of each type of whisky were differentiated from the competitor, helping consumers to identify the original product from the cheaper substitutes. Branding received further impetus following the industrial revolution and the evolution of commercialized societies during the eighteenth century. That was the period when city signs were introduced

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

on products as brand symbol. The method was adopted to indicate the origin of products, which in some cases had associations and perceptions related to the province or country of origin or to the producer itself. The current format of branding has come through different business concepts as companies sought long-term benefits. Luxury brands have an important role to play in the current corporate environment and they are a central part of business strategy.

Key Concepts for Luxury Brands

The luxury consumer. As consumers nowadays are looking for differentiation, they believe that luxury brands are part of the differentiation they are looking for. Consumers behave with a luxury attitude, which is portrayed when "luxury stimuli" is given to them. These luxury consumers are very much different from what they were in the past. They expect to be surprised and pampered, and are always pleased by surprises or when they get what they are asking for from the producer associated with their products. The association comes from the relationships that the producer builds with the consumer and from the way consumers are treated. Thus, we could recognize two segments of luxury consumers. The "traditional" and the "modern" consumer. The first segment is strongly attached to the traditional luxury brands as they knew them from their families. Brand tradition is very important for such families, and it guaranteed the avoidance of the risk of a new purchase. The modern consumers do not ask for brands or luxury products but for full support and service from the brands they buy. They are strongly aware of what they are looking for and are very much attracted by the experience that the brand gives rather than what actually the brand offers. This new generation has many different desires and characteristics compared to the luxury consumers of the past.

Thus, the relationship between the consumers and luxury brands has changed. Consumers are purchasing luxury brands because they think that these brand best fulfill their needs. Also, in some cases, these brands are the solution for a big problem in their everyday lives. Luxury brands help consumers to define what they really want, who they are, and how best they communicate amongst themselves and their peers. The strong attachment between the consumers and the luxury brands stems from the role of the luxury brands as symbols of personal identity. The modern luxury consumer offers his or her loyalty to luxury brands under the premise of being delivered the best possible, relying on its expectations. This expectation does not have a simple meaning, but has a dual character. It includes both functional and emotional benefits not only for the buyer but also for the users. In general, the symbolic needs are the most important ones as in some cases, consumers might find the functional ones in other brand types. Some of the intangible promises that consumers are getting are the ego of the consumer, the self-image, and self-esteem. The intangible level brings the consumer to a continuous movement from the rational to the irrational, which needs to be fulfilled. The main characteristics of the luxury consumer might be defined as follows:

- The luxury consumer is powerful. Consumers nowadays have more choices and more ways of purchasing the products they desire. Of course, the number of luxury brands has increased when compared to the past.
- The luxury consumer has high expectations and values. Luxury brands offer a variety of products and high standard of services. Consumers would like to understand the way these products are made and in many cases the material used to make the product. Consumers of luxury brands are persistent in their choices, they are strongly individualistic, but also socially disciplined, and they want to know the ethical issues related to the production of the luxury products they are buying.
- The luxury consumer is well traveled and well aware of the products and brand offerings. Luxury consumers nowadays are difficult to be cheated, in terms of product offerings. Consumers know, and are capable of analyzing, the brand or product messages and translate them to their benefit. They are also more informed through trips they undertake very often. They also know when to trade down in some categories of goods when purchasing luxury or other types of

goods. They also expect luxury brands to understand their needs, give them individual attention, interest and customization, which will give them more satisfaction from it. If a luxury brand is not satisfying them, they do not move on to another luxury brand, but engage in a dialog with the producer, building a relationship between the two. Such a relationship leads to strong brand loyalty, which is very important for the luxury brands as well as for the other brand categories.

As consumers have strong relationships with the brand producers, they need to have a continuous dialog between them. The following are the main principles of the dialog:

- luxury brands are synonymous with innovativeness and progress;
- luxury brands must embrace all types of products or services;
- luxury brands must have a clear value proposition related to the needs of current or future consumers;
- luxury brands must be adopted to the current environment;
- luxury brands must never negotiate their status.

Managing luxury brands. Brands are assets to the companies that own them. These assets come in an intangible form, and result in an added value for the company. In 2006, Interbrand placed a brand value of US$17.6 billion on Louis Vuitton, making it the most valuable brand in the luxury goods industry and among the 17 most valuable brands in any product category in the world. In addition to the benefits that these brands give to the companies, brands help consumers to identify the source of a product, reducing the risk in their purchases. Luxury brands are very beneficial to consumers because of the functional and emotional needs they satisfy. They can deliver valuable and exceptional promises and create strong brand perceptions. They can reduce the disappointment from other brand categories. The key words for luxury brands are innovativeness and differentiation. Differentiation is part of the tangible element of the brand, and innovation is interpreted differently by consumers. Thus, luxury brands combine both tangible and intangible aspects. Apart from these elements, in most of the cases, luxury brands have a strong emotional appeal.

How can the luxury brands be differentiated from other brand categories and which is the best way of doing this? The answer is – by enhancing the brand concept, and also by brand and product performance. The brand concept incorporates all the elements that could appeal to the consumers and the market. The brand concept is reflected through the brand name, the country of origin, the logo, and all the elements that are attached to the brand. Luxury fashion brands have a distinct brand concept differentiating them from others, although they share the prestige value. The horse carriage of Hermes logo represents the beginning of a very careful production that will last forever. The color is also unique to the brand and the taglines are in French reinforcing the country of origin of the product. The most visible aspect of a luxury fashion brand concept is the brand name. Usually, the brand names adopted are those of the founders of the company or of major designers of the products they sell.

Adopting a brand name for a luxury good requires careful handling from the marketing managers' point of view. The brand name is the tool that links the product with the market through different associations. Consumers need to easily identify the luxury brand among all the others. Under this consideration, some of the luxury brands, recognizing the symbolic value of the brand name, modify their brand names into more modern names that are more appealing to the new target groups or customers. A brand name for luxury goods needs universality. A luxury brand needs to be known and be remembered globally, even though the use of the product is rare, or the consumers are not aware of the usefulness of the product in their everyday lives.

Brand identity. Brand identity is one of the core elements of the brand and is strongly attached to the way consumers perceive the brand. Positioning is based on brand identity. The way consumers perceive the branded product has to do mostly with brand identity. If the brand

identity is clear then the success of the brand is guaranteed. Luxury brands have a very clear identity which relates the tangible and intangible aspects and the rational and emotional issues all together. The attributes and the identifiable elements of the brand constitute brand identity. The two major dimensions of brand identity are brand image and brand personality. Brand personality is the personality of the brand and has to do with the perceptions of the consumers or how the market perceives the brand. Brand image is the way the brand is seen by the people exposed to the brand communication. The differentiation between brand personality and brand image is sometimes difficult when speaking about luxury brands. Luxury fashion brands have a unique advantage in the brand personality and image development process because consumers perceive them as luxury. That is the reason luxury brands need slogans or characters commonly found in other consumer goods categories to accompany their brand symbols. Luxury brands share brand personality elements like originality classification, reliability and glamor. The levels to which the brands are adopting the above dimensions depend on the producer, the target group, and the market competition. In order to develop brand image and personality, the following five statements are crucial:

1. a holistic development with other brand elements related to the brand;
2. a clear positioning enhanced by the verbal and nonverbal communication the company adopts;
3. clear communication of the brand elements as a whole;
4. consistency over time and through ages and generations;
5. clarity in personality and the image communicated to the public.

Brand value. Brand value is the final outcome of the branding strategies. Brand value is the financial outcome for a company and leads to brand equity. Brand value is important for a company because in some cases it leads the brand being a very strong tangible asset, leading the company to higher levels of performance. Brand value is not a concept per se; it starts when the product is created and has an infinite end. For the time being, only a valuation method for luxury brands has been created. Companies that want to evaluate luxury brands are using methods created for other brand categories; these methodologies could not be considered appropriate for luxury brands.

Threats for Luxury Brands

Because of their high success, luxury brands meet counterfeiting. The four levels of counterfeiting that are very common nowadays include:

1. A counterfeit product which is the exact copy of the original product. Consumers believe that they are buying the real one or the genuine product.
2. An imitated product which is not identical but similar to the original one.
3. A craftsmen product which is similar to the original but the similarities are not so strong. In most of the cases the craftsmen producing these products have a connection with the brand.
4. A copied product, which is a copy of the original, and most often consumers are aware that it is fake. They are aware that the product is inferior in terms of materials used and way of production, but as they cannot buy the original, they prefer the copied product.

The luxury goods sector battles all the above fake products. These goods are marketed everywhere. These products damage the image of the original products, and, overall, the image of the brand. For the companies producing these products, there is a continuous fear of losing their prestige and facing a big failure of their marketing strategies. In 2003, counterfeit goods market cost New York more than one billion in taxes, an immense weight on the state's economy. Nowadays counterfeiting in fashion goods is very common. Counterfeiting was, and remains, a big problem for luxury industry for centuries. Luxury brands such as Burberry and Louis Vuitton are trying to battle counterfeiting by communication or by approaching the consumer and explaining the problems of using counterfeit

products. Louis Vuitton spends approximately €10–15 million per year for the battle. Part of this budget is given to lobbies, which try to influence governments to protect these products by banning their product or by strengthening their laws. Additionally, the company created an intellectual property bureau. It also employed agents who travel all over the globe in search of counterfeits factories and export orientation reporting them to the local authorities. The company also tries to educate consumers about counterfeited products. Louis Vuitton assists the seizing and destroying of counterfeit materials and finished products.

Managing the world of luxury business is quite fascinating. The development and implementation of effective branding strategies for luxury brands need constant revision and adaptability. Luxury brands have to be always innovative and creative in their marketing strategies. They should clearly deliver a strong message to their audiences. Every luxury brand needs to integrate different brand elements related to the market and to the consumers. There are techniques and tools that need to be used when planning and deciding about luxury brands. Some basic principles need to be adopted when luxury brands have strong and persistent presence in the market.

- Maintain a clear brand personality and identity
- Reinforce when necessary, especially when counterfeiting is prevalent
- Look for what competitors are doing and try to protect your brand
- Branding for luxury brands is quite important when consumers rely on it but also on the company
- Luxury brands are special brands which need special treatment.

See also *brand community; brand value; knowledge accessibility*

Bibliography

Aldridge, A. and Levine, K. (2001) *Surveying the Social World: Principles and Practice in Survey Research*, Open University Press, Buckingham.

(a) Baker, M.J. (2000) Writing a literature review, *Marketing Review*, **1** (2), 219–247; (b) Barzun, J. (1985) *Simple and Direct: A Rhetoric for Writers*, revised edn, University of Chicago Press, Chicago.

Becker, H.S. (1986) *Writing for Social Scientists*, University of Chicago Press, Chicago.

Belbin, R.M. (1993) *Team Roles at Work*, Butterworth Heinemann, Oxford.

Bell, R. (1992) *Impure Science: Fraud, Compromise, and Political Influence in Scientific Research*, John Wiley & Sons, New York.

Bell, J. (1999) *Doing Your Research Project: A Guide for First-time Researchers in Education and Social Science*, 3rd edn, Open University Press, Buckingham.

Bell, J. and Opie, C. (2002) *Learning from Research: Getting More From Your Data*, Open University Press, Buckingham.

Black, T.R. (1999) *Quantitative Research Design for the Social Sciences*, Sage, London.

(a) Blaikie, N. (1993) *Approaches to Social Enquiry*, Polity Press, Cambridge; (b) Blaikie, N. (2000) *Designing Social Research*, Polity Press, Cambridge; (c) Blaxter, L., Hughes, C. and Tight, M. (2001) How to research.

Block, P. (1981) Flawless consulting, Austin, TX: Learning Concepts.

Bogdan, R. and Taylor, S.J. (1984) *Introduction to Qualitative Research Methods: The Search for Meanings*, John Wiley & Sons, New York.

Booth, W.C., Columb, G.G. and William, J. (2003) *The Craft of Research*, 2nd edn, University of Chicago Press, Chicago.

Bryman, A. (1988) *Quantity and Quality in Social Research*, Routledge, London.

Bryman, A. and Bell, E. (2003) *Business Research Methods*, Oxford University Press, Oxford.

Bryman, A. and Burgess, R.G. (eds) (1994) *Analysing Qualitative Data*, Routledge, London.

Bryman, A. and Cramer, D. (2000) *Quantitative Data Analysis with SPSS Release 10 for Windows*, Routledge, London.

Buchanan, D., Boddy, D. and McCalman, J. (1988) Getting in, getting on, getting out and getting back, in *Doing Research in Organisations* (ed. A. Bryman), Routledge, London.

Buzan, A. (2000) *The Mind Map Book*, BBC Books, London.

(a) Cameron, S. (2001) *The MBA Handbook*, Financial Times/Prentice Hall, Harlow; (b) Campbell, J.P., Daft, R.L. and Hulin, C.L. (1982) *What to Study: Generating and Developing Research Questions*, Sage, Beverly Hills, CA.

Collis, J. and Hussey, R. (2003) *Business Research*, 2nd edn, Palgrave Macmillan, Basingstoke.

Cook, T.D. and Campbell, D. (1979) *Quasi-Experimentation: Design and Analysis Issues for Field Settings*, Houghton Mifflin, London.

Cook, C.K. (1985) *Line by Line: How to Edit Your Own Writing*, Houghton Mifflin, Boston.

Covey, S. (1986) *The Seven Habits of Highly Effective People*, Simon & Schuster, London.

Creswell, J.W. (1994) *Research Design: Qualitative and Quantitative Approaches*, Sage, Thousand Oaks, CA.

Crotty, M. (1998) *The Foundations of Social Research: Meaning and Perspective in the Research Process*, Sage, London.

Daft, R.L. (1984) Antecedents of significant and not-so-significant organizational research, in *Method and Analysis in organizational Research* (eds T.S. Bateman and G.R. Ferris), Reston Publishing, Reston, VA.

Davis, M.S. (1971) That's interesting! Towards a phenomenology of sociology and a sociology of phenomenology. *Philosophy of Social Science*, **1**, 309–344.

Davis, G.B. and Parker, C.A. (1997) Writing the doctoral dissertation: a systematic approach, Barron's Educational Series. Hauppage, NY.

customer relationship management and integrated marketing

Edward C. Malthouse and Bobby J. Calder

Customer relationship management, or CRM, includes all interfaces between consumers (or customers) and a company. It covers the interactions or contacts with the consumer, initiated either by the company or by the consumer (*see* CUSTOMER RELATIONSHIP MANAGEMENT). Hence, CRM activities occur when a customer calls a service number or when a current customer is sent an offer for a related product.

CRM is more than a name for these activities. *CRM* has come to be defined as an *informed approach* to interacting with consumers. CRM implies that the company should base its interaction with consumers on data. If a customer calls, the person answering the call should ideally have access to that customer's history of requests and their resolution. If the company attempts to cross-sell a consumer, the offer should be based on an analysis of that customer's previous purchases. CRM "systems" are thus designed to make interactions with consumers more data based (*see* PRODUCT DESIGN). The goal of CRM, as the name implies, is to turn interactions with consumers into ongoing relationships.

While the goal of CRM is to foster greater intimacy with the customer, there seems to be wide agreement that, in practice, CRM is often very tactical and impersonal. Often companies focus more on the hardware and software systems used in CRM than on the goal of making customer interactions more meaningful. This is even more the case in the United States than in other countries. There is even some indication that interest in CRM has recently cooled (as of the mid-2000s). Although CRM expenditures will no doubt grow in the long term, it increasingly seems that companies and vendors have been driven too much by a tactical IT approach.

From a longer term perspective, some companies are trying to make CRM "more strategic." As pointed out by Rigby and Reichheld, for example, companies have sought to define CRM as "the bundling of customer strategy and processes, supported by relevant software, for the purpose of improving customer loyalty and, eventually, corporate profitability" (Rigby, Reichheld, and Schefter, 2002). Note the emphasis on CRM as a way of turning strategy into activities.

Along these lines, Bligh and Turk relate CRM to Michael Porter's framework for strategy. Porter stresses that strategy can no longer be approached as mere operational excellence (OE) (Bligh and Turk, 2004). Such excellence is necessary but not sufficient to sustain competitive advantage. Strategy is about either cost leadership or activities that differentiate the firm from competitors. In Porter's terms, it needs to provide a competitive advantage in terms of cost leadership or marketplace differentiation. For example, CRM could be viewed as a way of reducing the total cost of dealing with consumers. By creating relationships with consumers, the company increases its efficiency in dealing with them. The line, however, between cost leadership and mere operational excellence is a fine one.

Thus, the most compelling way of approaching CRM more strategically is through seeking competitive advantage in the marketplace, though this has been relatively neglected in most applications of CRM to date. The big question is how can CRM contribute to competitive advantage in the marketplace? How can it *be* a competitive advantage? The answer is that companies need to connect CRM to the process of designing contacts for subsegments of consumers. The goal is to let these consumers experience the brand in a more individualistic way.

CRM and Subsegmentation

Segments are groups distinguished from others in a larger market. By definition, they are receptive to a brand in a way that other consumers are not. A segment may be defined in various ways (e.g., demographically, psychographically, attitudinally, or behaviorally), but it is always a group of consumers who currently – or potentially – value the brand in the same way. Segments share a common overarching receptivity to the brand (Calder and Malthouse, 2003).

On the other hand, *subsegments* are divisions of a basic segment into subgroups. By definition, a brand appeals to all of a market

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

segment. But the concept of subsegmentation goes beyond this. It implies that the brand can be further strengthened by building on the appeal of the brand for the entire segment by adding customized contacts. The idea is to customize interaction with the brand in some way for each subsegment. The marketer aims to allow different subsegments to experience the brand in a more individualistic or idiosyncratic way. These contacts are contingent on some unique characteristic of the subsegment's identity. One way of looking at this is that by customizing contacts with the brand to a specific subsegment, the brand can establish a relationship with the consumer.

CRM therefore becomes the key part of a marketing strategy. Relationship marketing is considered one of the highest forms of marketing strategy. Often it has no more substance than advertising that proclaims that the company wants to be the consumer's partner in something. But if implemented through CRM, contacts that really do depend on dealing with different consumers in a way that is contingent on who they are, the strategy can attain much more substance. At the same time, the brand itself provides leverage in that the brand appeals to the entire segment and is not being changed for each subsegment. This actually raises the possibility of both strategic differentiation and cost leadership.

In summary, the strategic use of CRM is to subsegment one's market in order to create contacts that allow specific subsegments of consumers to make a personal connection to the brand. At the same time, marketers can preserve the cost advantage of otherwise marketing to consumers in the same segment, the same way.

Bibliography

Bligh, P. and Turk, D. (2004) *CRM Unplugged: Releasing CRM's Strategic Value*, Wiley, New York.

Calder, B.J. and Malthouse, E.C. (2003) What is integrated marketing, in *Kellogg on Integrated Marketing* (eds D. Iacobucci and B. Calder), Wiley, New York, pp. 6–15.

Rigby, D., Reichheld, F., and Schefter, P. (2002) Avoid the four perils of CRM. *Harvard Business Review*, February, **32**, 101–109.

packaging and brand design

Bobby J. Calder and Steven DuPuis

Packaging plays a vital role in the advancement of brands. The package has evolved from being a way of safely transporting products to becoming one of the most important marketing tools for brands. Even the early Greeks marked their bottles with unique symbols to distinguish their maker. Modern day branding, as we know it, began in the early nineteenth century with the iconic packaging of Campbell's, Kellogg, and Quaker. Henry Parsons Cromwell developed the Quaker brand mark, becoming one of the first to create a cohesive marketing program. The packaging, advertising, and promotional elements were all consistently integrated and implemented.

Despite this history, in the recent past, marketers tended to underestimate the importance of packaging. It was considered one of the 4Ps and much less important than other areas of marketing: promotion, pricing, product, and distribution (place). Packaging was either at the end of the list of P's or a small part of the product P. Today, however, the importance of packaging is evermore recognized. It serves as the only form of communication that is seen by 100% of the brand's customers. In today's fragmented media environment, packaging is a critical form of marketing that uses strategic design to identify, inform, differentiate, persuade, and delight consumers to purchase. This has established an industry term, brand design, to describe this valuable marketing function.

What is Brand Design?

Branding is an effort to make products more meaningful to consumers. To do this one has to go beyond the product. One must transcend whatever the product is as a physical or objective entity. One must create and convey the meaning of the product so that the consumer's idea of the product is the idea that one wants them to believe. This is a process. It begins with an internal description of the meaning of a brand that is then transformed into something tangible that consumers perceive. If they perceive what the marketer intends, the result is a brand design. The result is a design in the sense that it fulfills its intended purpose.

The meaning can, of course, be expressed in many ways. It can be expressed verbally through words, names, and descriptive language. Or it can be expressed visually through pictures, symbols, colors, and shape/form. The goal of brand design is to use these verbal and visual devices to convey the meaning of the brand. In turn, consumers will perceive these as cues collectively and use them to make inferences about the brand.

Take a brand like UPS for example. At present, the central brand idea for this package delivery product is one of simplicity, functionality, and basic reliability. The UPS brand is something that one does not have to worry about. It is like a basic service: turn the switch, the lights come on. Send it UPS, the package arrives. The brand design aims to convey this brand idea in a way that the consumer perceives it. The design is based on the verbal and visual cues surrounding the product.

Brand design is different from the marketing communications task of telling the consumer about the brand with advertising messages. The brand design aims to let the consumer perceive the idea. Perception is fast and categorical. It proceeds and influences conscious thinking. The perception is more emotional than rational.

In the UPS brand design, the primary color is brown. What does the consumer perceive? Basic, down-to-earth, grounded. Just the perceptions intended by the brand idea. The design is successful.

The descriptive language features a slogan, "What can brown do for you?" The words are cues. Again the perceptions are routine, basic, and simple.

Visually the brand design uses a symbol in the shape of a brown shield – official, basic, and reliable.

The design, of course, spans uniforms, trucks, and packaging materials. It is also integrated into advertising in the form of an ordinary businessperson explaining how simple it is to use UPS by drawing a basic diagram with a brown marker on a whiteboard.

Brand design is thus the process of representing the brand idea in the reality of the consumer's world. All the consumer has to do is to take in the look and the feel of the product to

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

know what it is. Advertising can make use of it and build on it to tell the consumer more about the brand in an informational or emotional way. But the brand design speaks for itself and can be effective even without advertising.

The package of the 4Ps mindset is still there, but it is one of the presenters of the brand design.

Brand Design Versus Product Design

It is critical that marketers separate brand design from product design. The two are intertwined in the marketplace, but must be addressed differently. With product design, the marketer is concerned with how the product itself is physically experienced. This is likely to be measured in terms of product characteristics, for example, the sweetness of a drink, the sharpness of a picture, the stiffness of an automobile's ride. Most often, measurement is best accomplished with panels of people who are trained to detect and describe these product characteristics and who experience the product in environments that eliminate perceptual cues.

With brand design, however, marketers should try to minimize information about the physical product. It is best to evaluate brand designs with actual consumers. Since the goal is to evaluate perceptual impact, consumers should be exposed very briefly (a few seconds) to the design. It is important that they react in a perceptual mode – fast and automatically. More considered responses tend to obscure actual perceptual responses. One of the worst methods of brand design evaluation is to give consumers in a focus group (or in other types of open-ended interviews) a design and let them comment on it. Such settings are very likely to stimulate people to think about the design in all kinds of ways (including playing design expert) that obscure the design's real perceptual impact.

Even a good brand design can be undermined by a poor product design. In the marketplace, perception and the experience of the product itself come together. It may well be that a strong brand design can succeed even in the face of a weak product design. But this is by no means guaranteed. In the best case, brand design and product design are both optimized separately and come together synergistically in the marketplace.

Marketers and Designers

Marketers and designers often have different perspectives. Effective brand design requires more than just a division of labor, it requires collaboration too. Brand design is truly a process in which the marketer communicates the brand idea to the designer who translates it into reality. The designer makes the idea, so to speak, come to life. And the marketer evaluates the results based on whether the consumer's perception of the design matches the originating brand idea. Any lack of collaboration can cause a breakdown in the process at any point.

Design for marketing has evolved from its early beginnings as commercial art, visual communication, and graphic design to become a strategic methodology for brand design. This can be seen with the growing use of carefully planned visuals, such as mood boards that convey a brand's meaning, allowing marketers and designers to collectively communicate. Moreover, marketers and designers no longer think of this as a linear process. To some in the marketing community, design is still too often viewed as merely the implementation or execution of a brand strategy. However, in today's visual environment, design has become a vital element in the development of a brand strategy. We now see marketers and designers working together in the early stages of development, from innovation to brand positioning. A company's success is now seen through the collaborative efforts found within the diverse talents of marketers and designers.

Marketers and designers must stay connected in watching for societal changes that may affect consumer behavior. Strategies need to be reviewed and updated through collaborative efforts. Successful product launches and brand repositioning are achieved through strategies that begin with an integrated team approach. One would not enter a basketball game with the team members meeting for the first time on the court. Yet, in the past, we demanded success from designers that joined the team halfway through the marketing program. In the end, realizing and respecting the different talents that designers and marketers bring to the team is the key to collaborative success.

In some companies, design is being elevated as its role is seen as vital to the success and health of the organization. Many companies have begun to build design cultures internally, such as P&G, Apple, and Kraft, to name a few. They have made operational changes to include VPs of design. CEOs as well as marketers are increasingly seeing design as a key driver to increasing sales and ROI, and progressive companies are even giving design a seat in the boardroom.

These companies realize that an increase in competition and commoditization has led to product parity, leaving design and innovation as the only path to differentiation. As Daniel Pink says in his book, *A Whole New Mind*, "it is not enough that a product be reasonably priced and functional, it must also be beautiful, unique, and meaningful." Many retailers such as Target, Waitrose, Trader Joes, and Apple stores effectively use design as the primary way they influence consumers. Design plays to our emotions, inspiring rational thought by elevating our senses and creating a desire to interact. A study conducted by The University of Rochester School of Medicine found that eliminating emotions made it impossible for people to make decisions. Proof in the marketplace is Karim Rashid's sculpted trash can, Michael Grave's houseware line, and Deborah Adler's innovative packaging for prescription drugs, all successful, once-boring utilitarian items. When consumers are making a buying decision, they put themselves in the picture and evaluate the benefits and emotional risks of that decision. Even price is seen on an emotional level. Within a few seconds, buying decisions are rationalized on the basis of the facts, but the final choice is based on feelings. Humans buy when they trust and feel comfortable; they often justify the facts and buy on emotion. Reality is not perception, but perception is reality.

To stay competitive, companies must search for strategic advantages. Design is a key differentiator within a commodity-driven market place. Even as some companies continue to chase price and value, design and innovation plays an even bigger role in setting brands apart from their competition. Progressive companies look more and more to brand design.

Brand Systems

Closely related to the design of individual brands is the design of systems or portfolios of related brands (*see* BRAND EXTENSIONS AND FLANKER BRANDS). Brand systems must also be approached in terms of creating perceptions, but, in this case, these perceptions define a general core essence or similarity that holds together disparate products. Think of the Virgin brand umbrella, which covers everything from air travel to colas to electricity. Companies increasingly count on taking successful brands into new forms, alternatives, and categories. Thus Kraft has recently taken the Oreo brand from cookies into cakes. This is held together by the core perceptual essence of "fun to eat." David Aaker, among others, has written about managing such systems of brands (Aaker, 2004).

What permits a brand to become a system or portfolio of products is a brand idea that is broad enough to allow different product versions to be categorized as fitting the idea at a perceptual level. As with a single product, the perceptions expressed by the brand design can facilitate this.

Brand Design, Consumer Experiences, and 360 Degree Marketing

Increasingly, any brand must be an experience brand. Product parity and information overload is rampant. Without good brand design to inject life into the brand, this can be very difficult.

From Apple iPhone to Wal-Mart, brand designs serve as a holding device for our thoughts and feelings about things in our lives. They tell us stories that speak to our needs and wants. Their memorable mnemonics trigger cumulative memories and experiences. We look to them for inspiration, guidance, stability, and excitement. Each one comes with a promise and an expected brand experience.

Recently, we have seen a shift to fewer marketing dollars being spent on traditional advertising, largely due to the change in market conditions, media fragmentation, and an overcommunicated society. As a result, marketers place increasing importance on the in-store experience. A recent study by Deloitte Consulting found that 70% of purchasing decisions are made in-store and 68% of purchases

are impulse buys. This was important enough to have the former CEO Alan Lafley of P&G coin the term *first moment of truth* (the moment when consumers pick up a package). He built a successful corporate strategy around this consumer insight, called *shopper marketing*. This is not limited to just in-store merchandizing. Brand design must play a very significant role.

Retail environments and the products and brands that fill their shelves are seen by millions of consumers each day. Some brands win big in these settings while others lie unnoticed and neglected. This is the battlefield, the theater for brands to thrive or die.

In 1970, there were around 10 000 products on a grocery store shelf. Today, there are over 46 000. However, 50% of the products are not even seen by the average consumer and thus we are faced with an increased need for attention and impact. With all things equal, and with so many competitive choices, an effective brand design can be the tiebreaker in the purchase decision process. Within the retail environment, it is also the only place where brands will be seen with all their competition. So a brand design's effectiveness is critical. Brands and their packages must persuade, educate, and entertain in ways that delight and entice customers. Marketers and designers need to view this area as an opportunity to continually present their brands in new and fresh ways.

Ultimately, brand design speaks to the strategic creation of the customer experience. Each customer touch point or contact is critical to the overall acceptance of the brand idea. Each touch point that a consumer is exposed to defines and establishes his or her impression of a brand. Within these touch points, consumers will judge and assess whether they feel that the experience is authentic and worthy of their involvement. For many brands, brand design is the No.1 touch point with consumers.

There is also the need to keep pace with change. As the competition changes, so do the dynamics of the entire category. Brands are always on the move. As social and cultural views change, so does the visual language that represents these thoughts. The goal for marketers and designers is to respond in ways that keep their brands pure, fresh, authentic, and relevant. Marketers and designers need to look inward, be intuitive, and ask what has excited us and made us become involved with a product. After all, we are all consumers. Our ability to analyze social, cultural, and visual trends is critical; the skill is utilizing this information through collective collaboration and innovative hypotheses.

Established brands, of course, walk a fine line between embracing such change while not compromising their design equities and core brand idea. When Tropicana removed the orange with the straw from its package, a connection was broken with long-established design equity. Consumers found themselves cut off from their familiar perceptions of the brand. Furthermore, the redesign complicated the shop-ability of the brand, confusing customers further. Scott Young, President, Perception Research Services, states, "to avoid major mistakes, designers need to steer clear of brand hesitation and confusion at the shelf, thus ensuring shop-ability." The key to creating breakthrough at shelf is judging when and how to be disruptive with your brand's design equities.

The Future

In sum, we believe that "packaging" of the little P variety belongs to the past. The future is brand design. Marketers and designers must work together to bring brands from concept to reality in ways that express the meaning of the brand in verbal and visual ways that the consumer can immediately perceive. For many, if not most, products the marketer cannot wait for advertising to get the brand message across. The consumer must be able to know the brand simply by apprehending it.

Bibliography

Aaker, D.A. (2004) *Brand Portfolio Strategy*, Free Press, New York.

Grocery Manufacturers Association and Deloitte Consulting LLP (2008) 2008 Grocery Manufacturers Association (GMA)/Deloitte Consulting LLP Shopper Marketing Study, 13 Oct. 2008, 6 Dec. 2009 http://www.deloitte.com/view/en_US/us/Industries/ Consumer-Products/ article/4951af8ed510e110VgnVCM100000ba42f00aRCRD.htm

Pink, D.H. (2006) *A Whole New Mind*, Penguin Group, New York, p. 33

Young, S. (2009) Wins and losses at the shelf: what drives purchase intent? *Packaging World*, Nov., **58**, 34.

the concept of bundling: when the package contains more than its parts

Michael Kamins

Introduction

The term *bundle* is derived from the fourteenth-century English term *byndel* or *bindan* which means to bind together (Merriam-Webster, 2008). At that time, bundling was a practice in which a young man and young lady slept together in the same bed, usually fully clothed. What may seem strange today was common then as an acceptable form of interaction to allow parties of marriageable age to talk quietly when the day's work had been completed. They seemingly could relax together without the need to sit near the fire to keep warm—in effect they were all "bundled up". In marketing's use of the term, what is "bundled" are typically products and not people. The modern concept of bundling derives from the economics literature of the 1960s and the issues of tie-in sales where the firm chooses to sell the products together as a bundle and not separately (Guiltnan, 1987; Stremersch and Tellis, 2002). In an economic context, bundling has been studied from the perspective of a seller's motivation to bundle (e.g., is it an optimal strategy or not and under what conditions?), as well as consumer welfare and the impact of competition (Burstein, 1960; Stigler, 1961; Adams and Yellen, 1976; Telser, 1979; Schmalensee, 1984).

Types of Bundles

Price versus product. From its economic origins, bundling began to expand, developing into a relatively common and effective marketing strategy in which products were combined together into one offering. From the consumer perspective, bundling is based upon the idea that consumers value the grouped package more than the individual items that make up the package and appreciate the resulting simplification of the purchase decision that grouping provides. This strategy is very common, for example, in the software industry (e.g., consider Microsoft's Office which bundles a word processor, spreadsheet, email program, and presentation software), the cable industry (choice of basic and premium channels at one price), and the fast food industry (McDonald's Value meal includes an entrée, fries, and a drink for a discounted price). Bundling also offers various benefits for the seller inclusive of demand inducement and revenue enhancement. Whether they band together or operate as diversified entities in selling bundled goods, manufacturers benefit from improvements in costs via scale economies. Marketing firms and distributors that use different methods (like franchising, direct selling, and multilevel marketing) also stand to benefit through a faster turnover of inventory. Indeed, some industries can use bundling to effectively compete with foreign competition through the enhancement of product variety and value.

Stremersch and Tellis (2002, p. 56) define the term *bundling* in a marketing context as "the sale of two or more separate products in one package." These authors refer to both products and services. They further distinguish two different forms of bundling: that termed *price bundling* and a construct described as *product bundling*. In the former case, the separate products are sold at a discount without any attempt to integrate them into a cohesive whole. Such a form of bundling does not in itself create added value to the consumer and hence the maximum price consumers are willing to pay for the bundle (i.e., the "reservation" price), is by definition, equal to the sum of the conditional reservation prices of the separate products that make up the bundle. Therefore, in order to motivate purchase of the bundle, a price discount typically has to be offered. Examples of these types of bundles include purchase of Coca Cola and tickets to Great Adventure Theme Parks, or the bundling of Windows and Explorer by Microsoft, which was a focus of the government's antitrust case against Microsoft in 1998.

Product bundling on the other hand, involves the selling of separate products at any price. Here the combination of products into a bundle provides some segment of consumers with added benefits such that the consumer's reservation price for the entire bundle exceeds that of the individual elements that make up the bundle, and hence a premium could be charged. As Stremersch and Tellis (2002, p. 57) note, a

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

product bundle can therefore be thought of as having an integral architecture that relates to an underlying characteristic of the grouping (Ulrich and Eppinger, 1995). For example, the Optimum 3D network combines television, on-line and "voice" into one bundle providing the consumer with added convenience and efficiency. Price bundling has been used mainly as a promotional tool, whereas product bundling has been used more strategically since it provides the consumer, if done correctly, with added value.

Bundle Form. Bundles can also be characterized according to their form. One such dichotomy is termed *pure* versus *mixed* (Adams and Yellen, 1976; Stremersch and Tellis, 2002). When using a *pure bundling strategy* the firm is said to sell the products involved only in bundle form and not separately. As noted earlier, in the economics literature this strategy is also known as *tying* (Martin, 1999). It is this area of bundling that has received the most attention from the courts, as consumers have balked at having to purchase another additional item when their goal is to purchase a different item (Stigler, 1968). Clearly, this can be seen as problematic especially if the items contained in the bundle are not offered for sale separately and the consumer really does not want one of the items.

A "mixed" bundling strategy is used when a firm sells the items contained within the bundle separately as well as in bundle form. When this approach is used, it can have significant implications for the value that consumers put on the individual items when sold separately, specifically when one of the items is offered for free (Raghubir, 2004; Kamins, Folkes, and Fedorikhin, 2009). If a firm does not use a bundling strategy, one can describe their approach as "unbundling". However the use of such nomenclature only makes sense when one compares it to a bundling approach.

Bundle Content. The products that make up the bundle are also of interest from various perspectives. For example, research by Harlam *et al.* (1995) found evidence for the fact that bundles composed of complimentary items led to higher purchase intent than bundles of unrelated products. Interestingly, the authors received no support for their hypothesis that bundles composed of equally priced goods would have a higher purchase intent than bundles of unequally priced products. The authors believed that if a supplemental item of inconsequential value was included in the bundle, it would be largely ignored and the bundle would be perceived as essentially one product with little value added. Alternatively, if the prices of the items contained in the bundle are close to equal, both products' prices would be perceived to be relevant as key components of the bundle. However, in the limit, the impact of offering a free item in the bundle has an effect on the attributions that consumers make about the bundle as a whole (Kamins, Folkes, and Fedorikhin, 2009). Specifically, the inclusion of a free item could lead to the perception of value and hence a higher purchase intent, as opposed to the inclusion of an item that is not free but has a low price. Thus bundles can also differ as opposed to whether or not a "freebie" is included.

Finally, a distinction has been made regarding whether or not the items contained in the bundle are the same or different. The former case is known in marketing as *Buy One Get One Free* or (*BOGO*), while the latter is known simply as *multiproduct bundling* (Gaeth *et al.*, 1990). Recent research has combined the notion of "freebie" with the complexity of the bundle offered for sale (i.e., BOGO versus multiproduct bundle). Extending the work of Raghubir (2004), Kamins, Folkes, and Fedorikhin (2009) found that describing one of the disparate products in the bundle as "free" decreased the price consumers were willing to pay for each product when sold individually. In effect, giving something away for free (i.e., the supplemental item), led consumers to think of the free item in a negative light, that was reflected in a reduced perception of value when the item was sold individually. This finding replicated the results of Raghubir (2004). However, the "freebie" also had a detrimental effect on the focal (higher priced) item in the bundle when it was sold individually. This was explained by the fact that respondents attributed negative characteristics to the focal item (e.g., last season's offering, slightly flawed), if one had to give something away for free in a bundle to sell it. Interestingly, a "freebie" offer did not influence the

overall price for the bundle of disparate products while it did when the bundle contained identical products or BOGOs.

The differential effect of freebies on perceived value can be explained by the nature of judgment difficulty; that is, price inferences are easier to arrive at for just a single product than for a bundle, as well as for a BOGO than for a bundle of disparate products. Consistent with this explanation, factors that influence judgment difficulty (i.e., the salience of the company's motive for offering the freebie and time pressure to make a judgment) moderated the effects of a free offer on the amount consumers were willing to pay. This research implies that it may be detrimental to use "freebies" if one is using a mixed bundling strategy where the individual items contained within the bundle are sold separately. This is because the individual item may then be devalued. In addition, the seller should be cautious when selling BOGOs since the task is made easier for the consumer to derogate the single product that makes up the bundle.

Consumers' Perceptions of Bundles and Promotional Implications

A key question to ask in order to investigate this issue is "How are buyers' perceptions of value formed when presented with a promotion involving a bundle?" Specifically, when purchasers are presented with a bundle what strategy is best to use if you emphasize savings, and what is best to use when you emphasize costs? Kahneman and Tversky's prospect theory and Thaler's mental accounting theory are relevant to this issue (Kahneman and Tversky, 1979; Thaler, 1985). Prospect theory deals with how individuals make choices in situations where they have to decide between alternatives that involve risk. The theory proposes that individuals frame outcomes that differ from a reference point either as positive (gains) or negative (losses). Consumers are said to be much more sensitive to losses than they are to gains. Mental accounting theory attempts to describe the process whereby people code, categorize, and evaluate economic outcomes, such that multiple gains are perceived to be more rewarding and multiple losses are more repulsive than, respectively, a single gain (or loss) of the identical amount. This theory provides support for the "That's Not All Technique," discussed by Burger (1986) and utilized by numerous Television pitchmen inclusive of Ron Popeil and Billy Mays. For example, the statement:

"I'm not going to give you just this cookie cutter. Oh No! That's not all I'm going to give you. For the same price, I'm going to throw in a fine steel spatula. A bargain I hear you say? But wait . . . I'm going to make it even better, with this splendid temperature probe, absolutely free. Now, who wants this wonderful offer now all for only $9.99?"

Essentially, mental accounting theory suggests that the way a person subjectively frames a transaction in his/her mind will determine the utility he/she will receive or expect from the transaction. According to Thaler (1985, p. 202), "people try to frame outcomes in whatever way makes them happiest," and thinking about adding individual components into a deal for a fixed price is extremely pleasurable because of continual reveal of the additive nature of the deal.

These theories serve to help us determine the most effective way for the seller to present a bundle to the consumer. Consider the following example adapted from Yadav and Monroe (1993). Suppose a given seller can offer a "mixed bundle" of jewelry to the target market as illustrated in Table 1.

Using the example below, the question remains as to what is the consumers' perceived savings associated with buying the bundle? The way the information is presented suggests that there are two different types of savings for the buyer considering the purchase of the jewelry set. First, there are savings in relation to the purchase of the individual items ($359 – $269 + $199 – $149 = $140). Second, there is savings in relation to the purchase of the bundle ($418 – $319) = $99. So, how will the consumer frame the savings in relation to the purchase of the bundle? There are three potential alternatives:

1. The consumer compares the price of the bundle to the total sale price ($418 – $319) and arrives at $99.
2. The consumer considers the savings in relation to both the bundle ($418 – 319) = $99 and the individual items taken separately

Table 1 Components of a mixed bundle offer.

	Price If Items Are Bought Separately		
	Regular Price ($)	*Sale Price* ($)	*Savings* ($)
Item			
(A) Malachite necklace	359	269	90
(B) Malachite bracelet	199	149	50
Total	558	418	140
		OR	
Buy the Necklace and Bracelet as a set for $319, an additional savings of $99			

($359 – $269 + $199 – $149) = $90 + $50 = $140 for a grand total of $239.

3. The consumer simply compares the price of the bundle ($319) to the respective price of the items before they went on sale ($558) or ($558 – $319 = $239).

As individuals perceive multiple gains as more rewarding than a single gain, it was not surprising that Yadav and Monroe (1993) found evidence in support of the second frame above in terms of how consumers conceptualize a bundle. Thus when selling a bundle, it is critical for the vendor to clearly indicate the multiple savings that the bundle provides rather than the total savings, possibly by showing individual item discounts and the separate bundle discount. Alternatively, if the focus is on price and not savings, the seller should integrate all price information into a single bundle price rather than a list of separate product prices that make up the bundle, since the rehearsal of individual prices (since they are outlays) is painful. Finally, at some point, attention should be directed toward the relative savings of individual items within the bundle ($140 in this example) relative to the bundle savings ($99 in this example) and how the consumer's attributions about the relative size of these two different savings, maps out onto perceptions about the bundle and the items contained within it.

Promotional Implications and Future Directions in Bundle Research

One of the interesting characteristics of bundles is that when the seller uses the multiproduct form of bundling, which product serves as the anchor? This is an important question since if there is a great disparity in the price of the two items, will the higher priced item, in effect make the lesser price item seem to be of higher quality or will the reverse effect occur? The answer to this question, has important implications for the value of the bundle as a whole and hence the success of the promotion. While Gaeth *et al.* (1990) observed that consumers averaged the performance of separate items in the bundle to judge overall bundle quality (despite a large discrepancy in value between the two products), Drumwright (1992) found no evidence for this averaging process. This implies the dire need for more investigation on this issue.

An interesting perspective related to this issue has been offered by Kamins, Folkes, and Fedorikhin (2009). These authors suggest that a surprising combination of products (e.g., a bundle containing shampoo and a watermelon) is likely to elicit a significant amount of inferences about the bundle because disconfirmation of expectancies seems to stimulate attribution automatically (Wong and Weiner, 1981). This means that a seller's choice of items to bundle together is critical in the sense that discrepant value or characteristics can serve to motivate the consumer to disparage the bundle as well as the items that individually make up the bundle. As Harlam *et al.* (1995) found that items in a bundle, which serve as compliments to each other (as opposed to similar or unrelated products), result in a higher purchase intent.

Importantly, Kamins, Folkes, and Fedorikhin (2009) also found that the seller can manage the attributions that consumers make about bundles by providing favorable motives for why a bundle contains a freebie (e.g., as an anniversary celebration). However, for this approach to be effective, it is important to outline the conditions under

which attributions are elicited. For example, what is the impact on the joint effect of both price and nonprice information on bundle evaluation? When does one form of information take precedence over another and which leads to greater attributions?

Another area of fruitful research would be to examine the effect that various individual difference variables have on the consumer's perception of bundle value. One such variable, namely, the familiarity of subjects with the items in the bundle has been examined by Harlam *et al.* (1995). These authors found that more familiar subjects were more sensitive to price decreases (indicating a higher purchase intent) than less familiar subjects, but that no difference was observed between the two groups for price increases. Likewise, bundles and their various discounts require a significant amount of cognitive processing to be impactful, therefore, as proposed by Johnson, Hermann, and Bauer (1999), the consumer's "need for cognition" should serve as a fruitful individual difference variable when it comes to investigating the relative impact of bundling as a strategy.

Finally, we have seen above that the way in which bundle information is presented affects the frame that consumers use to evaluate the bundle and the items contained within the bundle. Hence there remains a need to examine how various methods of presenting bundle price information are framed, particularly in the case where an individual item contained within a bundle does not offer a price savings (Yadav and Monroe, 1993). Moreover, the bundle literature needs more focus on whether a price savings is more effective when assigned to the focal item as opposed to supplementary item within the bundle or whether this model is underspecified and other factors are more critical (see Janiszewski and Cunha, 2004).

Conclusion

Bundling is one of many different sales promotional devices available to sellers. As such, Raghubir, Inman, and Grande (2004) propose that its impact on consumers can be examined from three different perspectives: economic, informational, and affective.

According to these authors, the economic effect of promotion relates to monetary or nonmonetary (time and effort) gain or loss that a consumer promotion provides to the consumer. Thus the "cost" of consumption is viewed from a broader perspective than simply a monetary one. Hence in the case of a bundle, there can clearly be economic benefit, particularly, if the supplemental item is complementary and not necessarily offered for free. Indeed, one of the economic benefits outlined falls under the label of "convenience," which encompasses increased shopping efficiency due to reduced search costs. A bundle of a razor along with razor blades comes to mind.

The informational route communicates direct or informational knowledge derived from exposure to a promotion. In the bundle situation, there are informational benefits if the bundled promotion makes consumers aware of a product that better meets their needs.

However, for a product included as free within a bundle, consumers may discount its value when sold individually, simply because the approach signaled derogatory inferences about the item on behalf of the consumer.

Finally, the affective impact of promotion relate to the feelings and emotions aroused by exposure to the promotion, purchase of the promotion, or having missed the promotion.

For example, the purchase of a BOGO may make the consumer feel like a smart shopper while, at the same time, provide fodder for why the item was placed on "deal" in the first place.

In summation, the resulting impact of a promotion such as a bundle, cannot be simply measured by sales. At the very least, this article should suggest to the reader that many dimensions should be considered before a bundle promotion can be deemed as a failure or a success.

Bibliography

Adams, W.J. and Yellen, J.L. (1976) Commodity bundling and the burden of monopoly. *Quarterly Journal of Economics*, **90**, 475–498.

Burger, J.M. (1986) Increasing compliance by improving the deal: the that's not all technique. *Journal of Personality and Social Psychology*, **51** (2), 277–283.

Burstein, M.L. (1960) The economics of tie-in sales. *Review of Economics and Statistics*, **42**, 68–73.

Drumwright, M. (1992) A demonstration of anomalies in evaluations of bundles. *Marketing Letters*, **3** (4), 311–321.

Gaeth, G.J., Levin, I.P., Goutam, C., and Levin, A.M. (1990) Consumer evaluation of multi-product bundles: an information integration analysis. *Marketing Letters*, **2** (1), 47–57.

Guiltnan, J.P. (1987) The price bundling of services: a normative framework. *Journal of Marketing*, **51**, 74–85.

Harlam, B.A., Krishna, A., Lehmann, D.R., and Mela, C. (1995) Impact of bundle type, price framing and familiarity on purchase intention for the bundle. *Journal of Business Research*, **33**, 57–66.

Janiszewski, C. and Cunha, M. Jr. (2004) The influence of price discount framing on the evaluation of a product bundle. *Journal of Consumer Research*, **30**, 534–546.

Johnson, M.D., Hermann, A., and Bauer, H.H. (1999) The effects of price bundling on consumer evaluations of product offerings. *International Journal of Research in Marketing*, **16**, 129–142.

Kahneman, D. and Tversky, A. (1979) Prospect theory: an analysis of decision under risk. *Econometrica*, **47**, 263–291.

Kamins, M.A., Folkes, V.S., and Fedorikhin, A. (2009). Promotional bundles and consumers' price judgments: when the best things in life aren't free. *Journal of Consumer Research*, **36**, 660–670.

Martin, S. (1999) Strategic and welfare implications of bundling. *Economics Letters*, **62** (3), 371–376.

Merriam-Webster, Inc. (2008) *Merriam-Webster Collegiate Dictionary*, Springfield, MA.

Raghubir, P. (2004) Free gift with purchase: promoting or discounting the brand? *Journal of Consumer Psychology*, **14** (1&2), 181–185.

Raghubir, P., Inman, J.J., and Grande, H. (2004) The three faces of consumer promotions: economic, informative and affective. *California Management Review*, **46**, 1–19.

Schmalensee, R. (1984) Guassian demand and commodity building. *Journal of Business*, **57**, 211–230.

Stigler, G.J. (1961) The economics of information. *The Journal of Political Economy*, **69**, 213–225.

Stigler, G.J. (1968) United States v. Loews' Inc.: a note on block booking, in *The Supreme Court Review*, (ed. P.B. Kurland), University of Chicago press, Chicago, pp. 152–157.

Stremersch, S. and Tellis, G.J. (2002) Strategic bundling of products and prices: a new synthesis for marketing. *Journal of Marketing*, **66**, 55–72.

Telser, L.G. (1979) A theory of monopoly of complimentary goods. *Journal of Business*, **52**, 211–230.

Thaler, R. (1985) Mental accounting and consumer choice. *Marketing Science*, **4**, 199–214.

Ulrich, K.T. and Eppinger, S.D. (1995) *Product Design and Development*, McGraw-Hill, New York.

Wong, P.T.P. and Weiner, B. (1981) When people ask 'why' questions and the heuristics of attributional search. *Journal of Personality and Social Psychology*, **40** (5), 650–663.

Yadav, M.S. and Monroe, K.B. (1993) How buyers perceive savings in a bundle price: an examination of a bundle's transaction value. *Journal of Marketing Research*, **30**, 350–358.

sponsorship and event marketing

Pascale G. Quester

Introduction

Long considered as one of the many promotional tools available to marketers and brand managers, sponsorship has become one of the most flamboyant expressions of global branding in the later part of the twentieth century. Just as marketers have become aware of the power of associating their brands with key athletes or sports and cultural events, the marketing of such activities has also become highly sophisticated, turning events and other sponsorship properties into unique experiences that could sustain a branding of its own with their core audiences. As a result, sponsorship and event marketing have both developed in the last two to three decades into exceptionally professional and value-added marketing activities.

This article proposes to first define clearly what sponsorship and event marketing entail. It then offers a brief historical review of their development. The literature dealing with the impact of sponsorship and event marketing on consumers is reviewed next. The following section explores the issue of how to manage sponsorship and event marketing, an area which has received relatively little attention from scholars, leading to some strategies to defend sponsors and events' brands against the tactics used by ambush marketers.

The article then discusses the less apparent, but equally unique, opportunities afforded by sponsorship and event marketing and concludes with some predictions concerning the future of these two related but distinct areas of marketing expertise.

Definitions

Sponsorship refers to a marketing activity that involves investing in a property (e.g., an event, sport or art), with the view of deriving commercial benefits from the association, more particularly in terms of marketing communication, with this property. Sponsorship differs from philanthropy or corporate giving, in that the money invested is expected to deliver a positive financial return through brand exposure and, ultimately, sales.

Event marketing is the process by which a marketing strategy and a series of marketing activities are planned and implemented around an event, so as to ensure its profitability and commercial success. Event marketing typically involves securing corporate sponsorships as well as maximizing PR opportunities.

While modern commercial sponsorship is a relatively recent occurrence, with a dedicated special interest group of the American Marketing Association created in the mid-90s, sponsorship has existed in one form or another for a very long time. In ancient Rome, for example, it was common practice for returning provincial administrators or army generals to fund extravagant gladiator games to entertain the crowds in the hope of securing election to public office and lucrative official posts. In more recent history, kings and princes lavishly supported poets and musicians so as to cement their reputation as leading and cultured monarchs. So, while the intent of commercial sponsorship is clearly profit-driven, it is important to acknowledge its role in creating goodwill for the sponsors, be it with consumers, business partners, or employees of the firm.

The Growth of Sponsorship

Long confused with philanthropy or corporate giving, sponsorship has experienced over the last three decades a spectacular growth in terms of marketing expenditures dedicated to acquiring sponsorship rights. Over the same period of time, it has also become a topic of keen interest for academics and practitioners alike. In 2008, the year of the first ever Chinese Olympics Games, record levels of investments in sponsorship activity were recorded with total global sponsorship rights estimated at US$43.5 billion, an increase of almost 14% over 2007. In the Asian Pacific region, growth was larger in 2008, but from a lower base, with expenditures reaching US$9.5 billion, up 25% over 2007 (IEG Sponsorship Report, 2008). These figures, however, are but the tip of the sponsorship iceberg. Related marketing activities aimed at leveraging the investment by promoting the

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

association with target audience, such as advertising and PR surrounding the sponsored activity or event, are believed to represent three to five times as much as the value of the initial sponsorship right contract (Quester and Thompson, 2001).

In addition to its global and international reach, sponsorship offers a wide and diverse range of possible activities, from the small community-based grassroot sports programs to the largest and most dominant event in the sporting arena, the Olympic Games. Sponsorship can also be undertaken in a variety of contexts, although sports are by far the most popular area of investments, representing nearly 70% of all sponsorship budgets. This compares to only 3% of sponsorship budgets being allocated to arts sponsorship and 9% dedicated to a third party or cause, in what has been termed *cause-related marketing*. More difficult to circumvent and define, broadcast sponsorship, where sponsors invest money in order to "bring" to the audience a specific program, is a hybrid form between sponsorship and advertising, in that the message is as limited as it is in sponsorship but the exposure intensity and repetition are controlled and paid for, by the broadcast sponsor.

Several reasons can be put forward to explain the increased popularity of sponsorship with marketers. First, sponsorship provided a substitute method to achieve brand exposure for products and brands that are increasingly being banned from being advertised (such as tobacco products or alcoholic beverages). However, other products soon followed suit, as brand managers sought more exclusive platforms in order to avoid an increased degree of advertising clutter. In most developed economies, the progressive withdrawal of government funding, coupled with the increasing commercialization of sports and culture, created a context where corporate sponsorship has become a necessary substitute to government funding to deliver sleeker and more expensive entertainment to increasingly demanding audiences.

The conditions were set, therefore, for sponsorship to provide real potential as a communication and positioning medium for both local and global brands. Over the last two decades, sponsorship and event marketing have also grown in terms of sophistication and strategic significance to become an intrinsic building block for the equity of some brands, in what has been coined *sponsorship-linked marketing* (Cornwell and Weeks, 2005).

The Impact of Sponsorship on Consumers

Initially aimed at providing opportunities for brand exposure, sponsorship was first evaluated for its potential impact on consumers' awareness. Many studies have examined the impact of sponsorship on audience's recall and recognition of sponsor brands, revealing surprising levels of inaccuracy, which researchers have first established (Quester, 1997) and then sought to explain (Johar and Pham, 1999). For example, studies have shown that extreme levels of involvement into the event or activity sponsored prevented consumers from registering who the official sponsors were, so that moderate levels of involvement are preferable to either low or high levels of involvement (Walliser, 1993). Moreover, whether the sponsor is the market leader, whether the product carrying the sponsor's brands are related, or whether there was a congruence between the sponsor and the event itself also influenced whether consumers remembered the sponsor's brands accurately (Quester, 1997; Johar and Pham, 1999; Lardinois and Quester, 2001).

Later work has concentrated on evaluating the impact of sponsorship on other consumer responses, and more particularly attitudes. In particular, research has demonstrated that attitude change is predicated on the ability by the sponsor to leverage the association by widely communicating it (Thompson and Speed, 2000; Quester and Thompson, 2001). While there is no precise recipe on how much leverage should be applied to sponsorship, empirical research undertaken in the context of an art festival showed that a company investing one dollar for every dollar invested in acquiring sponsorship rights would derive substantially better results in term of attitudinal change than a sponsor who spent none, or less than one dollar (Quester and Thompson, 2001).

The impact of sponsorship on actual purchase behavior has not been fully established in the

literature, although some studies have examined the value of shares (that is, the probability that investors would buy shares of the sponsors), following sponsorship and naming rights announcements (Cornwell, Roy, and Steinard, 2001). Anecdotal evidence also suggests that sales of specific brands have benefited from prior involvement in congruent sponsorship investments.

Recognizing that awareness and attitudes are primarily cognitive processes, and that watching an event, sport, or art performance potentially engages consumers' emotions, researchers have recently suggested that emotions may play an important role in sponsorship effectiveness. These researchers have consequently called for any evaluation of sponsorship to monitor the affective response consumers have to the event, describing the persuasion process resulting from sponsorship as one of affective transfer (Christensen, 2006). Hence, as with advertising, sponsorship persuasion appears to follow a dual route of persuasion, via a cognitive path when the product is highly congruent with the activity involved (as is the case for engine oil sponsoring Formula One car racing) or via an affective route when the association allows the brand to capture the emotion of the event (as in the case for Coca-Cola's involvement in soccer).

However, while some aspects of sponsorship are comparable to advertising, the message remains eminently unscripted, in that no other message than the fact of the association is communicated to consumers. Hence, it befalls on consumers to make sense of this association and derive the meaning from it, a process assisted by the degree of fit or congruence that exists between the sponsor and the property it chooses to support. Although referred to in the literature under different other names, including fit and relatedness, the construct of congruence is the one that reflects both the degree to which the association is surprising to, or expected by, consumers as well as the degree to which it makes sense and allows them to derive meaning from the association (Fleck and Quester, 2007). For example, the sponsorship of a series of concerts at the Sydney Opera house by a pharmaceutical company would be considered unexpected for most consumers, until such time as they attended the event and received throat lozenges in the lobby of the opera house. The logical link between cough prevention and the need for a silent audience at a live music performance would then establish a clear and strong meaning which would assist consumers to recall the sponsor, and would lead to favorable attitudes toward that pharmaceutical brand. In cases where congruence is not so obvious, as when there is no such product-event use, or when it is less easy to determine by consumers, research shows that there is an advantage in articulating this to consumers (Cornwell *et al.*, 2006).

Clearly, property and event managers also have a stake in the way consumers respond (or not) to sponsors' unscripted message. For many of them, demonstrating the effectiveness of sponsorship in terms of consumer response is the key to securing further sponsorship agreements or the renewal of current ones. For event marketers, therefore, evaluating and demonstrating the type and magnitude of consumer response is at the core of developing a value proposition that "sells" this event to potential sponsors, as well as assist in the setting of the appropriate price for the sponsorship. Long deemed a mere transaction between two organizations, sponsorship has been more recently considered from a B2B perspective, whereby the sponsorship agreement involves the building of a long-term relationship between sponsor and property, with both brand managers and event marketers working toward mutually beneficial outcomes. This is described in greater detail in the next section.

Sponsorship and Brand Management

In a context where government and public funding has steadily decreased while both sports and art events have become increasingly commercialized, both sponsors and properties have become less transaction-orientated and more relationship driven. Indeed, the simultaneous efforts of product and brand managers and of event organizers have been such over the last decade or so that one can now describe sponsorship in terms of brand alliance (Farrelly and Quester, 2005) whereby the brand of the sponsor and that of the event are associated, for the mutual benefit of the two partners, in what can only be deemed a strategic long-term

marketing relationship. As such, much of the B2B research on relationship marketing should provide some perspective on how to manage sponsorship.

For example, research undertaken in the context of the Australian Football League has demonstrated empirically that the concepts of trust and commitment are just as important in sponsorship relationships as in any other type of B2B relationships. In particular, commitment, as measured by the degree of leveraging investment committed by the sponsor over and above the negotiated sponsorship fee, is a direct antecedent of its intention to renew the contract for a further term. Likewise, the degree of market orientation of both sponsor and event marketer was found to influence both the degree of trust and commitment sponsors associated with a particular sponsorship agreement. Hence, a longer term view, shared goals, and open and transparent communication were all factors that determined how effective a sponsorship was in the perception of a sponsor and in turn, how likely it was that this particular relationship would continue into the future (Farelly and Quester, 2003).

Sponsorship, however, can also be conceptualized as a much more complex form of cobranding, whereby "sponsorship moves from being a one-off exchange to being a long-term relationship between two or more organizations, and, as a consequence, sponsorship may be repositioned within the cobranded spectrum of the continuum" (Motion, Leitch, and Brodie, 2003). Indeed, research has established the relational nature of sponsorship and highlighted the need for sponsorship agreements to be market- and long-term orientated (Farrelly and Quester, 2005).

Corporate brands adopting such a cobranding approach to their sponsorship include Adidas and New Zealand Rugby Union (NZRU) (Motion, Leitch, and Brodie, 2003). Adidas became the successful contender for the NZRU sponsorship after they offered both money and also the opportunity to move from sponsorship toward building a powerful cobranded relationship that would expose the All Blacks to the global market (Motion, Leitch, and Brodie, 2003). In return, Adidas would gain media exposure and the chance to expand the size of the global rugby-apparel market (Motion, Leitch, and Brodie, 2003). Importantly, the brand values of the All Blacks were closely related to those of Adidas. Moreover, Adidas' corporate vision to create the best global sports brand reflected that of the NZRU to present the All Blacks as a leading international sports brand (Motion, Leitch, and Brodie, 2003). Such a long-term strategic commitment between a sponsor and a sponsored property results in a marketing strategy firmly built on long-term sponsorship.

Importantly, research shows that long-term, mutually trusting relationships between the sponsor and the property, the bedrock of any cobranding strategy, may provide some beneficial protection against ambush marketers and other firms seeking to derive benefits from misleadingly claiming an association with the same event. This is because genuine cobranding builds recognition by consumers of the long-term nature of the sponsorship, builds up the legitimacy of both partners, and, in effect, offers a buffer against any "pretenders" (Farrelly and Quester, 2005).

Cobranding provides access to the strategies of cobrand partners and the opportunity for each brand to pursue these strategies with the help and knowledge of an experienced partner brand (Motion, Leitch, and Brodie, 2003). As such, the opportunity to cobrand with a leading firm may represent one of the value-generating features of engaging in large-scale global sponsorship. Leading sports properties, including the International Olympic Committee or French Tennis Open Roland Garros, are purposefully reducing the stable of their sponsors, so as to enable them to work collaboratively over the long term. Furthermore, the literature has established that cobranding is a source of equity not only for product-specific brands, but also for corporate brands and that it provides a chance to dispose of inappropriate or negative values (Motion, Leitch, and Brodie, 2003). The investigation by Motion, Leitch, and Brodie 2003 demonstrated how two brands can move along the continuum from sponsorship to a successful cobranding relationship through the "construction of a unified identify ... that resulted in the formulation of successful linkages" (Motion, Leitch, and Brodie, 2003, p. 1091).

The Threat of Ambush Marketing

As budgets necessary to engage in sponsorship soared and as many of the prominent properties became locked in long-term relationships, many firms have found themselves unable to join in the club and have resorted to a strategy described as *ambush marketing*, whereby nonsponsors attempt to reap the benefit of an association without incurring the expense of negotiating an official agreement. Some of the methods these ambushers use involve saturating the media around the event, often using event-related themes, in order to create the impression that they, too, are in some way involved in an official capacity.

First identified by Sandler and Shani (1989), following the Olympics Games of 1988, ambush marketing refers to the attempt, by sponsors' competitors, to reap the same benefits as the official sponsors, without the preliminary investment in legally acquiring the rights to do so. Ambushers typically associate themselves with the event by appropriating cognate themes and saturating media with their brands in the immediate vicinity of the event (including before or after the television broadcast of the event, sport, or performance).

Sponsors and event marketers often attempt to prevent ambush by taking legal action, suing companies they believe are unduly exploiting an association with an event or activity to which they have not contributed and with which they are not, therefore, entitled to claim any legitimate link. However, legal protections appear relatively ineffective, as evidenced by the fact that during the 2008 Beijing Games, an ambusher, diary company Mingliu, achieved higher recognition scores than Coca-Cola, one of the top global sponsors for the event. Mingliu was successful simply by using saturation media buying, and using high-profile athletes as their brand endorsers. A better strategy would appear to involve the disclosure to consumers of ambushers' activities as research has shown this to not only erode the brand equity of the ambushers but also to bolster the brand legitimacy and credibility of the rightful sponsors (Mazodier, Quester and Chandon, 2010).

Other Benefits of Sponsorship

There are other aspects of sponsorship that appear to have been neglected. These are less frequently cited in the literature and yet they need to be empirically examined by sponsorship scholars. For example, while there is recognition in the literature that sponsors enjoy greater degrees of goodwill from consumers than do advertisers, there is little research exploring the potential of sponsorship to contribute to Corporate Social Responsibility profile, which in turn may also favorably influence both attitudes and purchase behavior on the part of consumers increasingly demanding ethical practices from business organizations.

The potential value of sponsorship in relation to internal marketing and human resource management has also been mentioned in the literature but remains unreported in the extant research, despite anecdotal evidence and case studies demonstrating that sponsorship can be successfully leveraged to bond staff and motivate sales people or service providers. For example, one of the greatest values derived by Credit Lyonnais, the French bank which sponsors the world famous cycling event Tour de France, is that the race stops at the end of every leg into a town or city where there is a Credit Lyonnais branch, giving the bank manager and their staff ample opportunity for hospitality, key client reward programs, and staff volunteering. At Credit Lyonnais, the summer months are dedicated to the running and following of the race, and this helps both in terms of staff recruitment and retention (Farrelly and Quester, 2005).

The Future of Sponsorship

The threat of ambush marketing, coupled with the global financial crisis of 2008–2009, has led experts to conjecture that 2008, the year of Beijing Olympics, may well have represented the high watermark in sponsorship expenditures, not the least because global properties are rare, but also because the layering of many levels of sponsorship (athlete, leagues, events, as well as the television broadcast of part of the event) can lead to such complex agreements as to induce the sort of clutter which sponsorship was supposed to avoid in the first place.

An additional threat is the recognized lack of control sponsors enjoy in relation to the image of the property they support. The chronic doping scandals surrounding the Tour de France, for example, are a predicament for the top sponsor of over 50 years, Credit Lyonnais. In the light of repeated instances of high level cheating, is a banking institution, which aims to project an image of trust and reliability, well advised to remain associated with such a high-profile event? At the same time, can it walk away from an event it has literally nurtured and been associated with for over 50 years, and which delivers such excitement to a large network of its bank branches across the national territory?

See also *attitudes; brand strategy; brand value; emotion; integrated marketing communication; integrated marketing communication strategy; persuasion*

Bibliography

Christensen, S.R. (2006) Measuring consumer reactions to sponsoring partnerships based upon emotional and attitudinal responses. *International Journal of Market Research*, **48** (1), 61–80.

Cornwell, T.B., Humphreys, M.S., Maguire, A.M. *et al*. (2006) Sponsorship-linked marketing: the role of articulation in memory. *Journal of Consumer Research*, **33** (3), 313–321.

Cornwell, T.B., Roy, D.P., and Steinard, E.A. II. (2001) Exploring managers' perceptions of the impact of sponsorship on brand equity. *Journal of Advertising*, **30** (2), 41–52.

Cornwell, T.B. and Weeks, C.S. (2005) Sponsorship-linked marketing: opening the black box. *Journal of Advertising*, **34** (2), 21–42.

Farrelly, F. and Quester, P. (2003) What Drives Renewal of sponsorship principal/agent relationships? *Journal of Advertising Research*, **43** (4), 353–360.

Farrelly, F.J. and Quester, P.G. (2005) Defending the co-branding benefits of sponsorship B2B partnerships: the case of ambush marketing. *Journal of Advertising Research*, **45** (3), 339–348.

Fleck, N. and Quester, P.G. (2007) Birds of a feather . . . definition, role and measure of congruence: the case of sponsorship. *Psychology and Marketing*, **24** (11), 975–1000.

IEG Sponsorship Report (2008) available from www.sponsorship.com/iegsr

Johar, G. and Pham, M.T. (1999) Relatedness, prominence, and constructive sponsor identification. *Journal of Marketing Research*, **36**, 299–312.

Lardinois, T. and Quester, P. (2001) Attitudinal effects of sponsorship on television audiences and the influence of sponsors' prominence: interaction and main effects of two types of sponsorship. *Journal of Advertising Research*, **41**, 148–158.

Mazodier, M., Quester, P. and Chandon, J.-L (2010), Unmasking the ambushers: Conceptual framework and empirical evidence, *European Journal of Marketing*, in press.

Motion, J., Leitch, S., and Brodie, R.J. (2003) Equity in corporate co-branding. The case of adidas and the all blacks. *European Journal of Marketing*, **37** (7/8), 1080–1094.

Quester, P.G. (1997) Awareness as a measure of sponsorship effectiveness: the Adelaide formula 1 Grand Prix and evidence of incidental ambush effects. *Journal of Marketing Communications*, **3** (1), 1–20.

Quester, P.G. and Thompson, B. (2001) Advertising and promotion leverage on arts sponsorship effectiveness. *Journal of Advertising Research*, **41**, 33–47.

Sandler, D.M. and Shani, D. (1989) Olympic sponsorship vs Ambush Marketing: who gets the gold? *Journal of Advertising Research*, **29** (4), 9–14.

Thompson, P. and Speed, R. (2000) Determinants of sports sponsorship response. *Journal of the Academy of Marketing Science*, **28** (2), 226–238.

Walliser, B. (1993) Modèle explicatif de l'influence du sponsoring sur le spectateur d'un événement sportif, in *Proceedings of the 9th Congress of the French Marketing Association (AFM)*, Association Française de Marketing, Marseille, pp. 668–697.

electronic public relations

Anastasios P. Panopoulos

Introduction

Internet as a communication medium overcomes the constraints of time and place, enabling the creation and maintenance of powerful relations between communicating parties. The lack of control, the decentralized structure, the open and public nature, the interactivity and independency between users are only few of the characteristics that contribute to the wide use of this medium. New online applications and functions, such as social networks, have brought revolutionary changes in the practice of public relations, creating innovative and more efficient ways in the invention, diffusion, storage and recovery of messages, thus altering the traditional communication patterns between organizations and stakeholder publics. Public relations practitioners are becoming increasingly involved in electronic public relations activities such as the launch of web sites, the monitoring of blogs and social networks, the support of e-voting activities, the creation of e-mails and e-newsletters, and so on. While in the past, organizations were identified as the source of information, controlling and directing the flow toward the stakeholder publics, a balance between the communicating parties has been established with the use of the Internet, since, in many occasions, the communication initiative now is undertaken by empowered Internet users (Ihator, 1999). Furthermore, the need for intermediaries or gatekeepers when trying to approach stakeholders is limited in an Internet context, in contrast to what is happening in the real world. According to Kent, Taylor, and White (2003) when it comes to electronic public relations the two-way symmetrical model can be seen as the process, dialogue as the product, and the Internet as the environment that makes it possible to happen. That is why we can argue that the two-way symmetrical communication model proposed by Grunig and Hunt (1984) has found its ideal communication channel since true dialogue between stakeholders can now take place. Furthermore, the Internet can be characterized as the first communication channel suitable for the transmission of both mass and personalized messages at the same time. For example, web sites as public relations tools can be used to disseminate symbolic content not only to large heterogeneous and geographically isolated audiences – mass communication – but they can also be used concurrently to build one-to-one relationships through the use of interactivity and customization – one-to-one communication. Web 2.0 technologies, in general, elevate the role of the Internet as a public relations medium, creating the need for a more effective strategic approach by public relations practitioners.

Paraphrazing Darwin's famous conclusion, public relations practitioners have to adapt to this new digital environment, as otherwise they face extinction. Traditional public relations activities are simply not enough, since a number of people choose to communicate only through the Internet. However, this does not mean in any way that the value and importance of off-line public relations is reduced. On the contrary, it stresses the need for a close synergy between actions taking place online and off-line. The two key elements for a successful public relations strategy in the new digital era are compatibility and continuation between traditional and electronic public relations. Public relations campaigns should be carefully redesigned to incorporate physical as well as digital elements. In this way, relations that already exist in the real world will be expanded and strengthened in the digital environment, while the overall effectiveness of public relations programs will be maximized.

The necessity for a successful integration of online and off-line public relations activities forces public relations practitioners to acquire new skills and competencies. Most public relations practitioners have no information technology background and, as a result, electronic communication is left to information technology departments with poor results. That is why public relations practitioners have to be transformed to empowered Internet users along with their professional public relations knowledge and expertise. In order for them to acquire some basic knowledge about information technology, training on internet applications and programming is required. Even when this is not the case, public relations practitioners should

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

always participate in the planning, designing, and monitoring of the digital environment leaving only the execution to the information technology departments.

The lack of continuous evaluation and the difficulty to define measurable goals for electronic public relations are two main obstacles that undermine their perceived effectiveness and their strategic role. In order to evaluate the real value and contribution of the Internet to the total outcome of public relations in an organization, specific measurable objectives aligned to the organizations' public relations strategy must have been set prior to the evaluation (Institute for Public Relations, 2001). Although different measurement techniques exist, the evaluation of online activities remains an unknown term for most public relations practitioners. Electronic public relations' effectiveness can be measured either at a public relations program level (preparation–implementation stages) or at a communication channel level (impact stage). A combination of the previous two levels in terms of economic value, web attendance, time allocation, requests of information, and so on, can reveal valuable information about online public relations activities. Various reasons, such as the lack of specialized personnel, resources, and methodology, are given by public relations practitioners to justify their aversion toward evaluation. However, none of the above is as significant as the results of the evaluation process itself.

There is no doubt that virtual worlds are becoming increasingly sophisticated, enabling organizations and stakeholders to actually step into and directly participate in a dialogic relational approach (Messinger *et al.*, 2009). Web 2.0 technologies enable a shift from the web as a passive information highway to the web as a dynamic platform for the exchange of real communication and experiences, since participation, self-expression, dialogue, and creation and maintenance of relationships within virtual communities is now allowed (Pfeil, Arjan, and Zaphiris, 2009; Wilson, 2006). Future developments will allow a further communication empowerment, enabling public relations practitioners to come even closer to larger groups of stakeholder publics in more effective ways. This article is organized as follows: electronic public relations principles are first described; in the next section, an analysis of effectiveness of electronic public relations is offered, and in the final section, the future of electronic public relations is discussed.

Electronic Public Relations Principles

The low cost for launching web public relations activities, the ease of implementation, the establishment of constant communication, and the attraction of a global audience are tempting arguments in favor of the adoption of electronic public relations for every organization. In addition to the aforementioned motives, the rush to secure an Internet presence along with the high competitive digital environment, has led organizations to adopt the Internet as a business panacea without first defining a clear strategic plan. However, web presence is simply not enough, since regular updating of the content of web pages and monitoring of the web environment are necessary actions, the hidden cost of which must be taken into consideration.

Broom, Casey, and Ritchey (1997) suggested that relationship formation and maintenance involve a process of mutual adaptation and contingent responses between parties. According to Thomlison (2000), *effective relationship management* is defined as the development, maintenance, growth, and nurturing of mutually beneficial relationships between organizations and their significant publics and, in order for these relationships to thrive, true dialogue must exist. The web incorporates text, sound, image, movement and the potential for real-time interaction, all in one package (Kent and Taylor, 2002). The element of real-time interaction differentiates the web from other communication channels such as television, newspapers, radio, and so on, where people behave in a passive way, without the ability to respond to messages or even to contribute to a feedback procedure. Through the commitment of organizational resources and training, the web can function dialogically rather than monologically (Kent and Taylor, 2002). As a result, two-way symmetrical communication is emerging as an important applied framework, as the practice of public relations moves toward a relational digital approach. Taylor, Kent, and

White (2001) managed to define the creation and maintenance of electronic relationships in the digital age, through the elevation of web sites as public relations tools. More specifically, they suggested that every web site, in order to promote dialogue and enhance relationships with the stakeholder publics should incorporate five dialogic principles coded as follows: usefulness of information, ease of use, conservation of users, generation of return users, and a dialogic loop. A sixth dialogic principle, coded as *web attractiveness* was added in the initial definition as shown below (Taylor, Kent, and White, 2001, p. 268):

> Dialogue first involves attraction (**web attractiveness**), whereby individuals or groups desire to interact (**usefulness of information**); for relationships to develop interactions must occur (**ease of use**); for relationships to grow dialogue must occur (**conservation of users**); and for relationships to thrive, maintenance and satisfactory interactions must occur (**generation of return users and dialogic loops**).

In this context, the public relations' dialogic principles can be expanded in order to be applied to any Internet application or activity performed by public relations practitioners. As a result, it is useful to analyze each dialogic principle under a public relations perspective (Taylor, Kent, and White, 2001).

Web Attractiveness. The beginning of every dialogic procedure stems from a certain attraction/need pushing both parties to communicate. It is the desire to exchange views with certain people that leads every one of us to enter a dialogic procedure in the physical world. Public relations practitioners have to boost that initial need in order to establish a continuous dialogue with all the stakeholder public. The projection of a digital corporate identity through the use of organizational elements such as the emblem/logo, the colors, photographs, and historic archives can improve the awareness of the organization's web activities attracting digital visitors. The combination of awareness and attractiveness in a digital environment is crucial since users are constantly bombarded with information/activities and they make decisions in just few seconds.

Usefulness of Information. The Internet has been named as the *information highway* since it offers quick and easy retrieval of any type of information. In fact, Internet users are bombarded everyday with information that they choose, voluntarily or even unconsciously, to absorb or to ignore. Online communications aim to convey information as well as relate to people. The usefulness of information is the criterion that is used by users in order to decide which information is of value. The essence of dialogue is the exchange of information between the participating parties; this information must be of interest for either side in order for the dialogic process to continue. Especially when a crisis is encountered, web activities, such as the update of the organization's web site, can provide continuous and timely information, attracting the interest of the stakeholder public.

Ease of use. Perceived ease of use can be seen as complementary to web attractiveness when it comes to first impressions. The use of the Internet for public relations purposes can be seen as an innovation, and, as a result, the reduction of technological complexity will help the adoption by the stakeholder publics. Efficiency, speed, and compatibility with the technological infrastructure available to the average user are crucial elements that enable the creation of user-friendly applications. Usability tests can provide a framework for the evaluation of web applications in order to secure a minimum level of ease of use that will enable the evolution of the dialogic procedure.

Conservation of users. The variety of options that every user has, combined with time restrictions of everyday life makes it hard to establish a long relationship in a digital environment. Usually, organizations and users are confined to the exchange of instant messages in an effort to communicate effectively, but in this way relationships can only be established while the maintenance and the quality of the relations is insecure. Relationship building and dialogue need time in order to flourish and prosper either in the physical environment or in the digital one. If organizations do not manage to provide a quality time framework when contacting their

stakeholder publics in a digital environment, the latter eventually will consider other choices.

Generation of return users. The principle of conservation of users should be aligned with the principle of generation of return users leading to the creation and maintenance of strong relationships. Users through repeated visits to older established web sites "prove" their digital relations with them. Public relations practitioners have to multiply the interactions with stakeholder publics in an effort to remind and cultivate relations with them. This is easier with the use of the Internet compared to other mediums, since it is possible to communicate with a low cost regardless of time or geographical barriers.

Dialogic loop. Interactivity and feedback procedures mark the passage from one-way communication to a fully dialogic one, enabling two-way symmetrical communication between organizations and stakeholder publics. Action from one of the two parties in the communication process and reaction from the other is a sign of dialogic behavior, creating a loop. Several applications like social networks, chat rooms, forums, and so on, support the creation of a dialogic loop attracting a high number of users.

Although public relations practitioners incorporate some elements related to the dialogic principles described above, through the use of technological applications, the lack of a dialogic orientation and their failure to apply a relational philosophy restrict the effectiveness of their digital actions. The main objective of electronic public relations is to create and maintain relations with different stakeholder publics in the digital world complementing activities and actions taking place at the same time in the physical world. As a result, every strategic plan adopted for electronic use should be aligned to the one created for the physical world, both projecting the overall public relations orientation of the organization. A dialogic philosophy could serve equally well both electronic and physical environments, allowing a continuous interaction through which a strong bond is created with the stakeholder publics.

Electronic Public Relations Effectiveness

The Institute for Public Relations (2001) recognizes the research function of effectiveness as every planned research action focusing in measuring, with the use of qualitative and quantitative variables, of the relative effectiveness of a program, strategy, or action, compared to predetermined goals. The evaluation of a web action/program should be constant and recurring in order to form a key success factor in an electronic public relations program (Cutlip, Center and Broom, 2000). It is obvious that the avoidance of evaluation is not an option. The measurement of effectiveness of a web action/program is a crucial function that can lead to the improvement of the web action/program and to the disclosure of potential errors, while, at the same time, it can be used as a continuous evaluation instrument. However, according to several studies, the initial approach of public relations' practitioners when designing web applications is confined to a trial–error basis without previous research. This approach can be extremely dangerous, since when stakeholder publics are not satisfied by the function or the content of a web action/program, they will easily turn to other Internet entities, rendering their return to their initial choice a difficult task. Results from usability studies reveal that stakeholder public can trace the information they are looking over the web in a relative small percentage even when they have access to the web page they want (Hallahan, 2001). As a result, evaluation and traffic analysis with the use of the Internet should be constant, extensive, tied with public relations objectives, and, above all, it should be conducted by the public relations departments.

Sometimes, even the web action/program evaluated can be used as a survey instrument through the creation of electronic questionnaires, the use of software metrics, or the performance of usability tests. Online web surveys are encouraged because of the low cost and the relatively high speed they offer. Today valuable online tools such as Google analytics (http://www.google.com/analytics/index.html), WebTrends analytics (http://www.webtrends.com), Shinystat analytics (http://www.shinystat.com), De-

epmetrix analytics (http://www.deepmetrix.com), and Awstats analytics (http://www.awstats.sourceforge.net), are available to public relations practitioners to help them gather information about their online activities. Information about the number of the web pages that a certain user viewed, the time spent on each, the web pages that used in order to come to and leave a web site, the user's perceptions and attitudes toward the web site, the archives that were downloaded, and so on. are only few of the valuable insights that these programs can offer to public relations practitioners.

The oldest and most frequent way of measuring the amount of traffic that an Internet application generates is through the number of hits that it concentrates (Holtz, 1999). Every visit to a web site is considered as a new hit, leading to the increase of the hit counter. This is an easy and simple way to monitor traffic in every web application without the active participation of the user. Furthermore, by recording the objects that the user chooses to hit, useful information about the function and user behavior toward the application would be revealed. However, there are certain reliability issues, since it is impossible to secure the uniqueness of the visitor or the hit. The reason for that is because each distinct visitor's action is recorded by the counter as a new hit increasing the total number of hits. For example, when a user is visiting the internal web pages of a web site or is coming back to the web site with the function of the back button, more than one hit is registered in the counter, creating a deceiving image of the number of visitors. This problem is even bigger if we take into account that in many organizations the web site is used as the home page or the initial web page when employees connect to the Internet.

As an alternative solution, the measurement of unique visitors in a web site has been proposed. This technique is based on the fact that a unique IP address is assigned to each computer in order to gain access to the Internet. As a result, if this unique IP address is recorded when a user is visiting a web site then we will know the exact number of the real unique visitors over a certain period of time. However, even with that approach, several issues of extreme importance for the public relations field remains unexplored. The number of visitors does not reveal any information about the usability of the web site, the information content, the conservation of visitors, or the capability to enhance dialogue through the web site. Separate measurement techniques can be used to address the previous variables like the monitoring of user behavior or the record of his/her actions. Finally, the association of electronic public relations objectives with economic values and the estimation of the bottom-line contribution to the overall economic result of the organization can also serve as a measurable indicator of effectiveness.

Unfortunately, many organizations focus only on recording instant impacts for a small period of time, forgetting that the main objective of the Internet as a public relations medium is not one-time communication, but the creation and maintenance of long-term dialogic relations. That is why multiple measurement techniques should be combined in an effort to evaluate and understand better the perceptions and attitudes of the users toward web sites (Holtz, 1999; Institute for Public Relations, 2001). Public relations practitioners are responsible for the determination of the variables that will be used in such combinations of techniques.

Future of Electronic Public Relations

Future developments in the field of electronic public relations will be determined by the evolution of information technology and the adoption of it by public relations practitioners. But in order for this to happen either in the physical or the digital world, time is needed. Already Web 2.0 technologies such as multi-user interaction, 3-D animation, open objective environments, and user-generated contents, audio and video transmission, blogging, instant messages, and chat have deeply affected the practice of public relations over the Internet giving a taste of what is going to follow. Internet users and as a result stakeholder publics are gaining more control day by day over their communication with organizations since new applications are even more user friendly and provide new capabilities. Social-networking web sites such as Facebook, MySpace, YouTube, and Linkedin, and games of virtual worlds

like Second Life have facilitated a greater freedom and realism to online communications. Public relations practitioners and organizations have made small steps in this direction by monitoring and creating profiles in an effort to exploit the power of electronic social networks in creating and maintaining relations. These online environs bring together most Web 2.0 technologies in simple, though highly usable, ways for people who have little to no technical expertise allowing them both content sharing and socializing (Messinger *et al*., 2009). Public relations oriented blogs and activities that build and maintain social capital (social support, integration, cohesion among stakeholders) in a Web 2.0 environment have promoted these sites as relationship-building tools (Dugan *et al*., 2008).

It is clear that the challenge for public relations practitioners will be to redesign their approach concerning relationship development with their stakeholders because of Web 2.0 applications. But before doing so, they must be empowered in terms of information technology in order to cope with the digital nature of their duties. Moreover, they have to find ways to secure a continuum and compatibility between public relations activities in the physical and in the digital world. Organizations and public relations practitioners that choose to ignore the electronic dimension of public relations as described earlier will have to struggle to survive in the new, rising, digital era.

Bibliography

Broom, G., Casey, S., and Ritchey, J. (1997) Toward a concept and theory of organization – public relationships. *Journal of Public Relations Research*, **9**, 83–98.

Cutlip, M., Center, H., and Broom, M. (2000) *Effective Public Relations*, 8th edn. Prentice-Hall, Englewood Cliffs.

Dugan, C., Geyer, W., Muller, M. *et al*. (2008) It's all "about you": diversity in on-line profiles. Proceedings of the ACM 2008 on Computer Supported Cooperative Networks, pp. 703–706.

Grunig, J. and Hunt, T. (1984) *Managing Public Relations*, Holt, Rinehart & Winston, New York.

Hallahan, K. (2001) Improving public relations s sites through usability research. *Public Relations Review*, **27**, 223–239.

Holtz, S. (1999) *Public Relations on the Net: Winning Strategies to Inform and Influence the Media, the Investment Community, the Government, the Public and More!* Amacom, New York.

Ihator, A. (1999) Society and corporate public relations – why the conflict? *Public Relation Quarterly*, **44**, 33–42.

Institute for Public Relations (2003) http://www.instituteforpr.com.

Kent, M. and Taylor, M. (2002) Toward a dialogic theory of public relations. *Public Relation Review*, **28**, 21–37.

Kent, M., Taylor, M., and White, W. (2003) The relationship between Web site design and organizational responsiveness to stakeholders. *Public Relation Review*, **29**, 63–77.

Messinger, P., Stroulia, E., Lyons, K. *et al*. (2009) Virtual worlds – past present and future: new directions in social computing. *Decision Support Systems*, **47**, 204–228.

Pfeil, U., Arjan, R., and Zaphiris, P. (2009) Age differences in online social networking – a study of user profiles and the social capital divide among teenagers and older users in MySpace. *Computers in Human Behavior*, **25**, 643–654.

Taylor, M., Kent, M., and White, W. (2001) How activist organizations are using the Internet to build relationships. *Public Relation Review*, **27**, 263–284.

Thomlison, T. (2000) An interpersonal primer with implications for public relations, in *Public Relations as Relationship Management* (eds J.A. Lendigham and S.D. Bruning), Lawrence Erlbaum Associates, Mahwah.

Wilson, J. (2006) 3G to Web 2.0? Can mobile telephony become architecture of participation? *Convergence*, **12**, 229–242.

attitude – behavior consistency

William E. Baker

Introduction

In this article, a brand attitude is defined as a "deliberately formed brand evaluation." It is both conscious and cognitive; it is not free-floating affect (Fazio, 1986). This definition assumes that attitudes are formed through a conscious cognitive process; they are products of deliberate expressions of summary evaluations. The source of the attitude may be "hot" affective responses (e.g., a strong emotional response) or "cold" information (e.g., quality cues, specific performance information). Besides distinguishing a brand attitude from a simple emotional response, this definition also recognizes and conforms to the methodological realities of attitude measurement. Whenever an attitude is measured, that measurement, by its nature, forces a cognitive component and deliberate expression, whether it existed prior to the measurement or not.

Brand attitudes play a central role in models of consumer behavior and models of advertising effects. Theoretically, often-studied and emulated approaches such as the classic hierarchy of effects (Lavidge and Steiner, 1961; McQuire, 1978), the Fishbein model (Fishbein, 1967), the Theory of Reasoned Action (Fishbein and Ajzen, 1975) and the Elaboration Likelihood Model (Petty and Cacioppo, 1981) focus on the process of changing or modifying brand (or object) attitudes. Methodologically, the vast majority of communication effects research relies on assessments of brand attitudes to test the persuasive impact of independent variables. More than 40 years after many attitude-based models of communication effects were developed, it remains fair to say that an explicit (i.e., theoretical) or implicit (i.e., methodological) assumption of most models of communication effects is that persuasion effects, and more specifically brand choice decisions, are mediated by brand attitudes. This presumption is reflected by the absence of research on the direct effects of information, marketing communications or otherwise, on brand choice. For example, 14% of 231 empirical advertising studies published in major marketing, advertising, and consumer behavior journals from 1990 to 1997 used brand choice as a dependent variable (McQuarrie, 1998). The rest either assumed the external validity of the brand attitude–brand choice relationship or were not concerned with it.

When studying attitude–behavior consistency in a marketing setting, it is important to recognize that there is a fundamental difference between the domain of social psychology, in which most attitude theory has been developed, and the domain of marketing. In social psychology, researchers are typically interested in studying people's attitudes and behaviors toward single issues, causes, or people. They are not, as in marketing, being asked to make a choice among several product alternatives. Also inherent in this distinction is the need to recognize that there are often significant differences between the environment in which attitudes are formed and the environment in which brand choices are made. These differences diminish the relevance of attitudes in choice contexts. Finally, it is questionable whether stable brand attitudes naturally exist in many product categories (Lastovicka and Bonfield, 1982). There are situations in which brand attitudes can be reliable predictors of brand choice; there are also many situations in which they cannot be expected to be reliable predictors. The purpose of this article is to discuss the factors that influence attitude–behavior consistency in marketing settings.

Do Stable Brand Attitudes Typically Exist?

Brand attitudes cannot mediate behavior unless they exist. Most communication effects modelers assume that their questionnaires measure stable brand attitudes in consumer memory. This, however, may not be the case. Lastovicka and Bonfield (1982) argue that people are likely to store firm attitudes toward social issues, but not typically toward branded products. As a result, responses to attribute belief and importance scales in survey research may often be misleading and reflect temporary tendencies, created on the spur of the moment, simply because a question was asked (see also Feldman

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

and Lynch, 1988). They further asserted that minor choice behavior can be best explained through principles of behavioral learning theory because low involvement decisions, which most brand purchases are thought to be, are unlikely to be governed by specific attribute beliefs.

According to functional theory (Katz, 1960), in order for attitudes to be held toward a given object, the object must be related to the concept of self, linked to an individual's important values, relevant to an important social group and perceived to give differential need satisfaction relative to other related objects. Using these criteria, it may be argued that consumers are unlikely to hold enduring attitudes toward brands in a great many product categories. The final criterion above adds particular insight into the problem of attitude development for many classes of branded products. It implies that consumers may have stable attitudes toward product categories (e.g., peanut butter, paper napkins), but not toward specific brands within categories (e.g., Skippy vs Peter Pan peanut butter). The lack of perceived interbrand differences in many product categories minimizes the potential for differential need satisfaction among brand alternatives within the product class (i.e., ketchup, paper towels, salt, etc.) and thus minimize the probability that brand attitudes can discriminate between alternatives (Baker, 2001).

People acquire denotative meaning about objects and issues, but evaluative and connotative meaning does not follow unless the brand carries significant personal importance. If one takes the position that most brands do not carry significant personal importance, it follows that consumers are likely to acquire knowledge of a brand that allows them to identify (e.g., "Crest is a brand of toothpaste") and describe it ("It's made by Proctor and Gamble." "It's the number one brand." "It has fluoride."), but not necessarily to evaluate ("It's an excellent brand." "It's the best brand.") or prefer it ("I'm going to buy it next time I need toothpaste."). Emotive and primitive evaluative responses, such as those postulated to be created by advertising, may result in purchase behavior, but do not necessarily lead to cognitively based attitudes. For example, repeated exposure to the brand name and package can lead to exposure-effect-induced purchase (Zajonc, 1980), and emotional associations "conditioned" to the brand by advertising can motivate buying (Baker, 1999), as can the operant effects of price discounting (Nord and Peter, 1980). All of these effects, however, neither require belief formation or attitude formation prior to purchase or after purchase.

Brand Attitude Accessibility

Assuming that brand attitudes are held in a given product category, the attitude must still be accessed before it can influence brand choice (Biehal and Chakravarti, 1986; Fazio, 1986; Kisielius and Sternthal, 1984). This is no trivial matter. People may not retrieve their attitude toward a stimulus before they act. There is no guarantee that there is anything but the weakest of links between the representation of the object and the representation of the attitude in memory (Bargh, 1984; Fazio, Powell, and Herr, 1983).

The likelihood that an advertising-generated brand attitude or any other attitude will mediate choice is a function of its salience at the time a decision is made. A primary mediator of attitude accessibility is repeated attitudinal expression (Fazio, 1986). In one series of experiments, attitude accessibility was demonstrated to be a direct function of the number of times subjects were asked to verbalize an existing object attitude (Fazio *et al.*, 1982). Accessibility was measured using a response latency measure. In subsequent experiments, Fazio found support for the potential of repeated attitudinal expression to lead to the spontaneous activation of attitudes when the attitude object is encountered (Fazio, Powell, and Williams, 1989).

However, experiments by Fazio and Herr (1984) demonstrated that even highly accessible attitudes do not necessarily directly or indirectly (e.g., bias consumer decision processes) mediate evaluation. In one experiment, subjects were told they were about to read a legal case concerning affirmative action. In a 2×2 design, half the subjects were asked to prepare to reach a verdict as they read the case; half were told to memorize all the factual details, but were not asked to reach a verdict. Also, half the subjects were asked to review their personal attitudes toward affirmative action prior to reading the

case, and half were not. Attitudes toward affirmative action correlated strongly to the verdict only in the cell in which prior attitudes were rehearsed and subjects were asked to prepare to reach a verdict. In the memory condition, attitude–behavior correlations were low regardless of attitude salience. Apparently, subjects found the specific information more relevant to their decision than their attitudes on affirmative action. The findings suggest that conscious or subconscious attitudinal mediation of behavior is not a necessary outcome of attitude salience during the decision process. This finding leads to the issue of attitude relevance (Baker, 2001) and diagnosticity (Feldman and Lynch, 1988).

Brand Attitude Relevance

Even if consumers hold brand attitudes in a product class and even if they are highly accessible in memory, there is still no guarantee that they will mediate choice at the brand response occasion. The question thus remains, when will attitudes formed at time A mediate choice at time B? Assuming that the conditions above are met, evidence indicates that the answer is when the attitude formed at time A is more relevant, that is, more diagnostic than other salient information at time B.

Much of the available research on the effects of prior judgments and attribute information on evaluation or choice is not designed in such a manner that it can specifically address the above question. A large body of evidence demonstrates that memory-based judgments can dominate the evaluative implications of recalled attribute information on subsequent evaluation (Dreben, Fiske, and Hastie, 1979; Lichtenstein and Srull, 1985; Lingle and Ostrom, 1979). An equally large body of evidence reveals a dominance of attribute information over previously formed judgments (Biehal and Chakravarti, 1986; Reyes, Thompson, and Bower, 1980). The research successfully demonstrates effects of judgment and effects of attribute information, but does not specify clear boundary conditions for these effects other than accessibility and does not offer a theoretical framework to explain or predict when each effect is likely to occur.

Feldman and Lynch (1988) built a framework that is able to reconcile much of the conflicting evidence in the literature. In the context of brand choice, Feldman and Lynch propose that the likelihood that any cognition about a brand, whether generated by marketing communications, personal experience, or word of mouth, will be used to choose that brand is a function of three factors: the accessibility of the input in memory, the accessibility of alternative brand cognitions in memory, and the relative diagnosticities of the inputs.

Brand attitudes that are the output of deliberately formed summary evaluations (i.e., cognitive elaboration) are likely to have an accessibility advantage in memory over more specific brand information because they are likely to be more efficiently stored in brand memory (Alba and Hutchinson, 1987; Wright, 1980). For this reason, they are more likely to become the product of automatic processing (Bargh, 1984) and less likely to become the victim of output interference (Alba and Chattopadhyay, 1985; Hoch, 1984). A simple judgment referral heuristic, that is, the use of an attitude to guide choice behavior, is also a convenient and simple way to make decisions (Bettman, Johnson, and Payne, 1991). Ceteris paribus, this gives attitudes an advantage over other more specific information. In Lichtenstein and Srull (1985), both attribute information and prior judgments were given an equal chance of mediating evaluation. In this scenario, judgments formed at time A were accessed and used in favor of the attribute information at time B. However, there was no reason for subjects to make the effort to use attribute information in their evaluation at time B. The evaluation environment was exactly the same at time A and time B; there were no brands, no new brand information, no changes in the purported usage occasion, and so on. The likelihood that a previously formed judgment will be at least as relevant and at least as diagnostic as other accessible information is a function of the similarity between the context in which the attitude was formed and the context in which the choice is made. In reality, unless attitudes are formed at the time of choice, there will almost always be differences.

A Change in Context between Attitude Formation and Brand Choice

The assertion that brand attitudes measured at time A can reliably predict behavior at time B requires that two key assumptions be accepted. First, one must assume that the attitude formation process and the brand choice process are identical. The pool of information and the rules used to transform this information into evaluations must be identical for the outcomes to be reliably identical. Second, one must assume that the pool of available information, the accessibility and relevance of this information, and the motivation to use that information are unchanged from time A to time B. Evidence suggests that the validity of these two assumptions is highly questionable (Baker and Lutz, 1988).

Bettman (1979) demonstrated that many identified choice-based strategies such as lexicographic, conjunctive, and elimination by aspects do not require an overall evaluation of brand alternatives during the choice process, but rather only evaluation or recognition of specific performance attributes or quality cues. Other research affirms the position that different informational sets and evaluation strategies appear to drive attitude formation and choice decisions (Johnson and Russo, 1984; Bettman and Zins, 1977). If this is true, there is no reason to trust the ability of attitude measures to predict choice outcomes.

If there is a change in the evaluation environment between the time of attitude formation and the time of subsequent evaluation, evidence suggests that specific performance information is more "diagnostic" than prior summary evaluations and, hence, will be used in favor of prior judgments. Consumers perceive prior judgments to be nondiagnostic if new brands, attributes, or usage situations are introduced into subsequent evaluation contexts (Biehal and Chakravarti, 1986; Lynch and Srull, 1982). Experiments designed to test Lynch and Feldman's basic propositions have supported the framework (Lynch, Marmorstein, and Weigold, 1988). In one experiment, subjects were put in "mixed choice" tasks where they were forced to choose between a previously evaluated brand and a new brand. Subjects relied on attribute information, and not the prior judgment, to make the new evaluation if the attribute information was accessible, ostensibly because the prior judgment had no diagnostic implications in the new evaluation setting.

Another reason choice and judgment processes are likely to differ is a decision involvement gap. One antecedent of decision-making involvement is the perceived risk associated with the purchase (Robertson, 1976). There is little risk associated with forming a "wrong" attitude when asked because there are no real consequences associated with the evaluation. In many product classes, however, there is considerable economic, social, and utilitarian risk associated to a "wrong" choice. This involvement gap between attitude and choice formation processes is likely to lead to evaluations based on different informational inputs and different decision rules. When this occurs, attitude–behavior consistency should be low. For this reason, survey measurement procedures that simulate the true "state of mind" (e.g., involvement) at the time of choice are more predictive of choice (Baker and Lutz, 2000).

Assuming equal levels of involvement at the time of attitude formation and choice and even identical evaluation processes, the correlation between brand attitude measures at time A and brand behavior at time B is still likely to be low unless involvement is high (Baker, 2001). Momentarily activated cues have a disproportionate influence over judgments made about an object or related behaviors performed shortly after their activation (Wyer and Srull, 1986; Kisielius and Sternthal, 1984). Environmental influences have a tremendous effect on salience. Cognitions that influence evaluation are cued to memory by contextual factors and only a small subset of the available information will be accessed at any one evaluation episode (Feldman and Lynch, 1988). Since accessibility is a function of the time since the most recent activation of that cognition (Wyer and Srull, 1986), attitudes measured at or near the time of advertising exposure are likely to be heavily biased by advertising content. Attitudes measured in another environment are likely to be driven by the most salient information in that environment.

Attitude–behavior consistency is likely to be particularly low when consumers engage in low-involvement decision-making behavior. In such a scenario, the latitude of acceptance

of alternative courses of action is likely to be high (Sherif, Sherif, and Nebergall, 1965) and evaluations are likely to be a function of highly salient contextual cues. Assuming that attitude measurement immediately follows exposure to advertising, the salient contextual cues are likely to be advertising effects. Assuming that choice takes place in a store, the salient contextual cues are likely to be packages and other point-of-purchase (POP) information for both the "target" brand and competing brands. When involvement is high, attitude–behavior consistency should be higher because evaluations are more dependent on deliberately accessed rather than cued information.

Meta-analytic Evidence for Attitude-behavior Consistency

A series of meta-analyses have been conducted in recent years, which supports the theoretical arguments and empirical findings discussed above (Glasman and Albarracin, 2006; Kraus, 1995; Cooke and Sheeran, 2004). An examination of hundreds of studies indicates that attitudes correlate more strongly with behavior when they are (i) accessible and stable, (ii) frequently repeated, (iii) formed on the basis of direct experience with the attitude object, (iv) formed on the basis of one-sided rather than two-sided arguments, (v) formed in contexts informationally similar to the choice environment, and (vi) formed in environment characterized by high involvement. If one or more of these conditions do not exist, attitudes do not tend to reliably explain behavior. Certainly, the reliability of attitudes formed vicariously in contexts disparate informationally, motivationally, and in time from brand choice environment must be considered questionable in terms of their predictive validity.

Summary

There are several barriers to using brand attitudes, advertising generated or otherwise, as reliable predictors of brand choice. First, brand attitudes may not be held toward brands in many product categories. Second, brand attitudes may be neither stable over time nor accessible at the time of brand choice. Third, a salient brand attitude must be perceived to be at least as relevant as other salient information at the brand response occasion. Vicariously formed brand attitudes rarely possess such diagnosticity. Fourth, the attitude and brand choice process must be identical. Evidence suggests that this is not typically the case. Informational inputs and evaluation rules are likely to differ. Fifth, unless involvement is high at the time of attitude formation, the effects of varying environmental cues and time delays on information accessibility is likely to lead to different informational inputs to the brand attitude formation process and the choice process. The Feldman and Lynch (1988) framework reconciles much of the conflicting research on the role of prior judgments in a choice context. Its stated conditions for an effect of prior judgments on choice, accessibility, and diagnosticity, have stood the test of time.

In many respects, research on attitude–behavior consistency suggests that if brand choice is the ultimate focus of interest, then brand choice itself should be the dependent variable that is studied and tracked. If actual purchase behavior cannot be studied in a given situation and measures such as brand attitude or purchase intention must act as surrogates, then steps should be taken to match the research environment as closely as possible (i.e., information, motivation, competitive alternatives, reason for purchase, etc.) with the brand choice environment.

Bibliography

Alba, J.W. and Chattopadhyay, A. (1985) The effects of context and part-category cues on the recall of competing brands. *Journal of Marketing Research*, **22**, 340–349.

Alba, J.W. and Hutchinson, J.W. (1987) Dimensions of consumer expertise. *Journal of Consumer Research*, **14**, 411–454.

Baker, W.E. (1999) When can affective conditioning and mere exposure directly influence brand choice. *Journal of Advertising*, **28**, 31–46.

Baker, W.E. (2001) The diagnosticity of advertising generated brand attitudes in brand choice contexts. *Journal of Consumer Psychology*, **11** (2), 129–139.

Baker, W.E. and Lutz, R.J. (1988) The relevance-accessibility model of advertising effectiveness, in *Nonverbal Communication in Advertising* (eds S. Hecker and D.W. Stewart.), Lexington Books, Lexington, pp. 59–84.

Baker, W.E. and Lutz, R.J. (2000) An empirical test of an updated relevance-accessibility model of advertising effectiveness. *Journal of Advertising*, **29**, 1–14.

Bargh, J.A. (1984) Automatic and conscious processing of social information, in *Handbook of Cognition and Social Learning* (eds R.M. Sorrentino and E.T. Higgins), Guilford Press, New York, pp. 1–43.

Bettman, J.R. (1979) *An Information Processing Theory of Consumer Choice*, Addison-Wesley, Reading.

Bettman, J.R., Johnson, E.J., and Payne, J.W. (1991) Consumer decision making, in *Handbook of Consumer Behavior* (eds T.S. Robertson and H.H. Kassarjian), Prentice-Hall, Englewood Cliffs, pp. 50–84.

Bettman, J.R. and Zins, M.A. (1977) Constructive processes in consumer choice. *Journal of Consumer Research*, **4**, 75–87.

Biehal, G. and Chakravarti, D. (1986) Consumers use of memory and external information on choice: macro and micro perspectives. *Journal of Consumer Research*, **12**, 383–405.

Cooke, R. and Sheeran, P. (2004) Moderation of cognition-intention and cognition-behaviour relations: a meta-analysis of properties of variables from the theory of planned behaviour. *British Journal of Social Psychology*, **43** (2), 159–186.

Dreben, E., Fiske, S., and Hastie, R. (1979) The independence of item and evaluation information. *Journal of Personality and Social Psychology*, **37** (10), 1758–1768.

Fazio, R.H. (1986) How do attitudes guide behavior, in *Handbook of Motivation and Cognition* (eds R.M. Sorrentino and E.T. Higgins), Guilford Press, New York, pp. 204–243.

Fazio, R.H., Chen, J.-M., McDonel, E.C., and Sherman, S.J. (1982) Attitude accessibility, attitude-behavior consistency, and the strength of object-evaluation association. *Journal of Experimental Social Psychology*, **18**, 339–357.

Fazio, R.H. and Herr, P.M. (1984) Attitude accessibility following a self-perception process. *Journal of Personality and Social Psychology*, **47** (2), 277–286.

Fazio, R.H., Powell, M.C., and Herr, P.M. (1983) Toward a process model of the attitude-behavior relation: accessing one's attitude upon mere observation of the attitude object. *Journal of Personality and Social Psychology*, **44**, 723–735.

Fazio, R.H., Powell, M.C., and Williams, C.J. (1989) The role of attitude accessibility in the attitude-to-behavior process. *Journal of Consumer Research*, **16**, 280–288.

Feldman, J.M. and Lynch, J.G. Jr. (1988) Self-generated validity and other effects of measurement on belief, attitude, intention and behavior. *Journal of Applied Psychology*, **73**, 421–535.

Fishbein, M.A. (1967) Attitude and the prediction of behavior, in *Readings in Attitude Theory and Measurement*, (ed. M. Fishbein), John Wiley & Sons, New York, pp. 477–492.

Fishbein, M. and Ajzen, I. (1975) *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*, Addison-Wesley Publishing Company, Reading.

Glasman, L.R. and Albarracin, D. (2006) *Psychological Bulletin*, **132** (5), 778–822.

Hoch, S.J. (1984) Availability and interference in predictive judgment. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **10**, 649–662.

Johnson, E.J. and Russo, J.E. (1984) Product familiarity and learning new information. *Journal of Consumer Research*, **11** (1), 542–550.

Katz, D. (1960) The functional approach to the study of attitudes. *Public Opinion Quarterly*, **24** (2), 163–204.

Kisielius, J. and Sternthal, B. (1984) Examining the vividness controversy: an availability-valence interpretation. *Journal of Consumer Research*, **12**, 418–431.

Kraus, S.J. (1995) Attitudes and the prediction of behavior: a meta-analysis of the empirical literature. *Personality and Social Psychology*, **21** (1), 58–75.

Petty, R.E. and Cacioppo, J.T. (1981) *Attitudes and Persuasion: Classic and Contemporary Approaches*, William C. Brown Co, Dubuque IA.

Kim, M.-S. and Hunter, J.E. (1993) Attitude-behavior relations: a meta-analysis of attitudinal relevance and topic. *Journal of Communications*, **43** (1), 101–142.

Lastovicka, J.L. and Bonfield, E.H. (1982) Do consumers have brand attitudes? *Journal of Economics Psychology*, **2** (1), 57–75.

Lavidge, R.J. and Steiner, G.A. (1961) A model of predictive measurement of advertising effectiveness. *Journal of Marketing*, **25** (2), 59–62.

Lichtenstein, M. and Srull, T. (1985) Conceptual and methodological issues in examining the relationship between consumer memory and judgment, in *Psychological Processes and Advertising Effects: Theory Research and Application*, (eds L.F. Alwitt and A.A. Mitchell), Lawrence Erlbaum, Hillsdale.

Lingle, J. and Ostrom, T. (1979) Retrieval selectivity in memory-based impression judgments. *Journal of Personality and Social Psychology*, **37** (2), 180–194.

Lynch, J.G. Jr, Marmorstein, H., and Weigold, M.F. (1988) Choices from sets including remembered brands: use of recalled attributes and prior overall evaluations. *Journal of Consumer Research*, **15**, 169–184.

Lynch, J.G. and Srull, T.K. (1982) Memory and attentional factors in consumer choice: concepts and research methods. *Journal of Consumer Research*, **9** (1), 18–37.

McQuire, W. (1978) An information processing model of advertising effectiveness, in *Behavioral and Management Science in Marketing* (eds H.L. Davis and A.J. Silk), Ronald Press: New York, New York.

McQuarrie, E.F. (1998) Have laboratory experiments become detached from advertiser goals? A meta-analysis. *Journal of Advertising Research*, **38** (6), 15–24.

Nord, W.R. and Peter, J.P. (1980) A behavior modification perspective on marketing. *Journal of Marketing*, **44**, 36–47.

Robertson, T.S. (1976) Low commitment consumer behavior. *Journal of Advertising Research*, **16** (2), 19–24.

Reyes, R.M., Thompson, W.C., and Bower, G.H. (1980) Judgmental biases resulting from differing availabilities of arguments. *Journal of Personality and Social Psychology*, **39** (1), 2–12.

Sherif, C.W., Sherif, M., and Nebergall, R.E. (1965) *Attitude and Attitude Change*, Saunders, Philadelphia.

Wright, P.L. (1980) Message-evoked thoughts: persuasion research using thought verbalizations. *Journal of Consumer Research*, **7**, 151–175.

Wyer, R.S. Jr and Srull, T. (1986) Human cognition in its social context. *Psychological Review*, **93**, 322–359.

Zajonc, R.B. (1980) Feeling and thinking preferences need no inferences. *American Psychologist*, **35**, 151–175.

the role of creativity

George G. Panigyrakis and Prokopis K. Theodoridis

Introduction

One of the most important components of an integrated marketing communications program (*see* INTEGRATED MARKETING COMMUNICATION) is the advertising message. As advertising is a form of communication, it involves a message that is communicated by the sender, usually the company, to the receiver, who is the consumer. The objective for the company is that consumers interpret the messages as the company intends. From everyday life, it is easy for everyone to realize there are various ways to convey an advertising message. However, it is the creative strategy that decides whether the advertising message is well presented and executed (Belch and Belch, 2009).

Good creative strategy and implementation both play an essential role in the success of a product or a service (*see* BRAND STRATEGY) and can diminish the progress of an already established brand. Conversely, an advertising campaign that is poorly implemented can result in the product being a failure. Creativity in advertising is one method applied by agencies to affect persuasion, and, eventually, behavior. Due to its purpose, advertising creativity is different from art creativity, because it must accomplish goals decided by other players. This is not generally an occurrence in the arts (Till and Baack, 2005).

Many ads have won awards for creativity but these ads failed to increase sales (*see* ADVERTISING EFFECTIVENESS), resulting in advertising and marketing people becoming critical. They argue that creative people in ad agencies are sometimes more interested in creating award-winning ads, than successful ads, which accomplish their clients' objectives. However, other advertising people believe awards are a good way to recognize creativity that often does result in effectiveness.

A successful creative idea is manifest. It is etched in the mind of the consumer and can have an enormous impact on sales. This is also consequential in the appointment or dismissal of advertising agencies. Agencies that win creative awards secure more business and can therefore afford to retain better staff (West *et al*., 2008).

Creativity is the foremost but least scientific aspect of advertising. It must possess impact, quality, style, and substance. Ideas must be original and moreover pertinent to the product and target consumer to facilitate beneficial solutions to marketing communications problems. "Creativity is the ability to produce work that is novel (i.e., original, unexpected)." Novelty is a required but inadequate prerequisite for creativity as it must also possess adaptability and appropriateness. (Sternberg and Lubart, 1996, p. 3). The combination of "novelty" and "appropriateness" or "usefulness" is generally accepted as taken. These notions are explored in the following sections.

Creativity in Advertising

Creativity has been defined by every synonym of the word creativity, plus, value added, and acceptable. However, the most used descriptor of creativity in advertising appears to be *novelty* – a divergence from the norm – and a sense of uniqueness or originality. (Ang and Low, 2000). The foundation of creativity in advertising, similar to other disciplines, is knowledge. It has patterns, rules, conventions, language, and symbols. Creative ideas are novel. They are thoughts and/or images that are generally recognized but considered foreign to one another. Leo Burnett, when talking about creativity, supports this statement. When talking about creativity he did not use the term original. Instead, he referred to making meaningful associations between previously unrelated thoughts in ways that were relevant, believable, and in good taste (Burnett, 1968). Other researchers prefer the word original. (El-Murad and West, 2004). Original and novel appear to be analogous. However, the dissimilarity is of importance. Novelty is the quality of being different, new, or strange. Originality, conversely, is the state or quality of being newly created or formed, fresh. (Oxford Advanced Learner's Dictionary, 2009). Consequently, creativity in advertising, while appearing to be original, is in fact novel.

Although divergence is a central determinant of creativity, the ad must also be relevant. The

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

relevance component of creativity reflects the extent to which ad elements are meaningful, useful, or valuable to the consumer (Smith *et al.*, 2007).

While the consumer seeks relevance, practitioners, mostly regard creativity as business objective. Practitioners seem to be rather more in harmony with what their clients require and not necessarily with consumers. This may explain why creativity and its effect remain unpredictable Consumers receive advertising messages according to their needs. Their opinion of creative can contrast from the presumption of creatives (West *et al.*, 2008). It is accepted that consumers and creatives disagree on creativity, but, there is insufficient evidence as to why. It has been attested that creative ads containing novelty, meaningfulness, and emotional content were perceived as more favorable, resulting in a more favorable view of the brand itself and increased purchase intent. It was also found that an ad considered to be creative is one that is unexpected, carries a pertinent message, and generates positive feelings (Ang and Low, 2000; Till and Baack, 2005).

"Creative" advertising and "effective" advertising are the concepts (*see* ADVERTISING EFFECTIVENESS) that most frequently appear in the literature (Ogilvy, 1985). However, even though creativity in advertising has been widely recognized as very important, the link between creativity and advertising effectiveness has not been extensively examined, leading to diverse opinions on the role and importance of creativity in advertising and marketing (Till and Baack, 2005).

In the business environment managers tend to value "effectiveness," which is usually measured by changes in awareness levels, consumer attitudes, or in market sales, whereas creative people generally have a low regard for these types of measure. There are people who believe that advertising is creative only if it sells the product or services, and that the impact that advertising has on stimulating consumers' purchase intention is more important than its novelty and originality. In contrast, others evaluate creativity in terms of innovative, inventive, and aesthetic values. They believe that creative ads can capture the consumers' attention and influence decision making (Belch and Belch, 2009). Perspectives on advertising creativity depend on one's role, leading to many different views on how advertising creativity is perceived.

Even among trained advertising professionals, there is a lack of agreement, resulting in a struggle between those who create and those who manage. It has been also mentioned that even copywriters and art directors may have differing views. Part of the problem can be traced to what the various parties involved consider as priority. Clients are more concerned with strategic questions of appropriateness rather than issues of novelty. Product managers and account executives consider ads as promotional tools and evaluate them according to their impact on their clients' objectives. Moreover, when referring to strategy and sales, appropriateness for them could also mean effectiveness, time, and money, as well as corporate values. Conversely, the art directors and copywriters consider ads as a medium to express and communicate their own points of view and their personal objectives (Hirschman, 1989).

Ultimately, it is the combination of these viewpoints that constitutes creativity in advertising and makes an ad effective and successful. Besides being original and interesting, an ad must make a good impression on consumers, trigger their attention and, influence their decision, as well as being original and interesting.

In the next section, a paradigm of creativity in advertising is presented through a systems approach incorporating all potential "actors" and influences on creativity.

A Paradigm of Creativity in Advertising

The creative process is not just one person with "the idea." There are several other interrelational sectors and players to consider. They can be placed into four categories: The field, the domain, the creators and the external environment (Csikszentmihalyi, 1999), all parts of which have an influence on the whole.

The field workers: the account handler, research department, accounts services, accounts planning, creative director and clients brand managers, category managers, marketing directors, and a chief executive officer, and

included occasionally are the audience – the customers.

The domain: the product, competitors, marketing communication.

The creators: the writers, the arts department, and the creative director.

The external environment: photographers, film directors, costume designers, and all other relevant media people.

It all begins with an "advertising brief," given to the agency's account handler by the client. This brief consists of a few standard parameters including the general message the client wishes to communicate, details of the product/service, and selling points. The client will also suggest which medium is to be used, that is, TV, press, radio, or direct marketing, and inform the agency of the maximum budget. The client's main contact in the agency is the "account handler," whose job it is to turn an ad brief into a creative brief. This is done by using the services of all departments in the "field," especially research, who examine the "domain."

To facilitate the creative process, many agencies now use account planning. Planners work with the client as well as agency personnel, to discuss how information they have gathered can be used in the development of the creative strategy.

The account handler then presents the final brief to the creatives. His brief contains information on the target market, the message, and the media requested by the client.

The creatives then work on proposals. A creative team consists of an art director and a copywriter; a mixture of pictures and words.

Now the proposed ad has to be priced. This is where production, also known as *traffic*, enters. It is their job to take into consideration the countless factors and elements that go into physically producing the ad. They deal with all players including the "external environment."

The creative director is not only responsible for the quality of the final creative work but also has the task of creating harmony amongst all the players. This can be a very difficult task as tempers can rise (for example, a creatives idea is blocked by accounts) or the client, through the account handler, insists on a location or actor not included in the budget. They are often praised highly when their team's efforts win awards, but conversely, the creative director shoulders the negativity when a project goes wrong or response falls short of expectations.

Creative Strategy – Approaches and Appeals

Planning creative strategy is a great challenge. For those who work on producing the ads, it is very difficult to create an ad that is unique, innovative and at the same time effectively communicates the advertising message. Ads must be novel, unusual, meaningful, and relevant to the consumers' interests and needs. It is often difficult for the creative team to accomplish such a challenge and create an effective and successful advertisement. Advertisement, as every marketing situation, is different and each campaign requires a different creative approach. Every client has his/her own specific goals and objectives that need to be achieved. Many creatives believe that clients must take some risks, if they want to create an ad that will break through the clutter and grab consumers' attention, but not all companies or agencies have the same opinion. Many managers argue that being risky is not always effective.

In the context of winning creative ideas, risk-taking perspectives may limit the number of creative ideas produced and offered. An agency working for a client who is antirisk will need to limit creator imagination. Simultaneously, an agency that is antirisk will behave in the same manner. Therefore, risk is necessary to develop original creativity (Belch and Belch, 2009).

The literature supports the fact that the client will have the most influence on the advertising agency's risk taking. It is often the case that for smaller clients, advertising agencies are likely to propose greater risk taking. This implies that given the consequence of potential failure, agencies are likely to attempt to protect themselves from negative impact (El-murad and West, 2003).

The popular belief is that creativity in advertising is best considered as a process. An organized approach to problem solving to develop a creative advertisement is a safe way for success. James Webb Young (2003), developed one of the

most popular approaches of the creative process, containing five steps:

Immersion: gathering information through background research.

Digestion: thinking about the information and working on them.

Incubation: putting the problems into perspective.

Illumination: coming up with a great idea.

Reality or verification: refining the big idea.

There is however, a six step plan derived from the previous approach. These steps are used by the majority of advertising agencies presently and are: Discovering, Strategizing, Creating, Evaluating, implementing and finally Measuring.

These six steps plus the full creative jargon are now presented.

Discovering. This involves the preparation and collection of useful background information and is the initial step in the creative process. It is very important to not only understand and identify the nature of the company and the product but also, those of competitors. Actual customers, target markets, economic environment, social trends, and new developments in the field must be taken into consideration. This research is indispensable for complete comprehension. Agencies must possess savoir-faire to synthesize the information they gather in order to develop an advertisement. To facilitate the process, many agencies now use account planning which is a process focused on clients wants and needs. One must first know the product in order to create an effective advertisement. Quantitative and qualitative research facilitates this procedure (focus groups, ethnographic research) facilitates this procedure enormously (Belch and Belch, 2009).

Strategizing. This is the analysis of all the discovery facts to find interesting relationships that lead to the exploration of creative procedures. The decision is then made to either focus on a single strategy or an interteam task where ideas are considered. The task is then passed on to the creative department.

Creating. Selecting the campaign, organizing a work plan, and finding the big idea of the campaign are taken into consideration at this stage. For an ad to be successful, it must contain a unique idea that captures the consumers' attention. The ad also has to appeal to create a positive reaction toward the product. Furthermore, a creative ad has to ensure that the product or service is differentiated from competitors' offers.

It is quite difficult to put all these together in order to create a successful ad, but there are some approaches that can assist the creative teams in developing an effective advertisement. These include *using a unique selling proposition* (USP), *creating a brand image, finding the inherent drama, and create positioning.*

As Rosser Reeves (1961) stated, a USP has three major characteristics. First, each advertisement must make a proposition to the consumer, not just promoting selling. Second, this proposition must be original and attractive for the consumer, and third, it must be strong enough to motivate new customers toward the brand.

Creating a brand image is also very important in creative strategy. The image and the personality of the brand should be strong, easy to memorize, and have an appeal to the potential users. It must also differentiates the product from similar ones.

Another approach is *finding the inherent drama*, or the special characteristics and benefits of the product or service that make the consumers want to purchase it.

Finally, *positioning* holds a great role on determining the major selling idea, because it establishes the product in a specific place in the consumers mind.

Different advertising appeals are used to intrigue consumers' attention, attract their interests, and influence their feelings toward the product. They motivate the consumers to have a positive reaction. There are several categories/approaches of advertising appeals.

Informational/rational appeals focus on the logical view of the consumers. Their intention is to promote the benefits of the product or service that satisfy consumers' realistic needs. They emphasize on the practical effects of the product, such as quality and/or, efficiency that are important to consumers and increase their purchase intentions.

Emotional appeals focus on exciting customers' feelings. They are designed to influence consumers on an emotional level and affect their

social and psychological needs. Many consumers are thought to follow their feelings to make a decision and pay less attention to the rational benefits that a product or a service offers.

There is also a combination of *rational and emotional appeals* that are used to appeal to consumers that make their decisions based on both emotional and practical motives, to purchase a product.

There are also additional types of appeals that differ from rational and emotional appeals, and can be classified as *reminder advertising* and *teaser advertising*. Reminder advertising focuses on building brand awareness and teaser advertising focuses on building curiosity and excitement about a product or a brand.

Advertising execution relates to the way advertising appeal is used to communicate the advertising message and how it is presented. An advertising message can be executed in various ways (Belch and Belch, 2009):

Animation: a recent and also popular approach, usually for commercials targeting children (creative, entertaining ads).

Combinations: combinations of two or more execution techniques.

Comparison: a direct way of communicating a brand's particular advantage over its competitors.

Demonstration: illustrating the key advantages of the product/service by showing it in actual use.

Dramatization: focusing on telling a story with the product or service as the star.

Humor: a type of appeal but also a way to express/present other appeals.

Imagery: the ad consists primarily of visual elements such as pictures and illustrations an symbols rather than information.

Personality symbol: developing a central character or personality symbol that can deliver the advertising message and with which the product or service can be identified.

Scientific/technical evidence: cite technical information, such as the results of scientific studies to support the advertising claims.

Slice of life: on the basis of a problem/solution approach, portraying a problem that consumers might face in their daily lives, with the ad showing how the product can resolve it.

Straight sell or factual message: a straightforward presentation of information concerning the product or service.

Testimonial: using a person to praise the product on the bases of his/her personal experience with it.

Priken (2004) provides more ideas in ad execution.

Execution styles – the kickstart catalogue.

Absurd, surreal, bizarre: associating the brand with absurd or bizarre ideas, representing the benefit of the brand with a surreal or fantastic situation.

Alternative media: using an outdoor site or installment, trying to find unconventional ad formats, using familiar places or objects in a provocative way.

Alternative uses: finding new use of the brand into a context or environment, and revealing a new perspective.

Break out of frame: using the advertising medium or its context or an altering version of it into the advertising.

Change of perspective: showing objects or situations from unusual viewpoints that is, extreme close-up or distance, or take on other identities.

Change the product: the way of presenting the product differently to the audience changing its shape, use, location, packaging, whatever can be changed.

Come and play: involving the audience in a game.

Comparative juxtaposition: or "before and after" – making the benefit of the brand obvious by a comparison.

Double meanings: visual (optical illusions) or verbal (word play or suggestion), involving double meanings.

Exaggeration: description of features of the brand, problem situations, or solutions that can catch audience attention and emphasize the benefits.

Metaphor and analogy: comparing the benefit of the brand with something similar.

Mixing and matching: combination of two concepts or objects to produce something new.

Omission and suggestion: missing or incomplete information, replacing with something that emphasizes its USP.

Paradoxes: a contradictory statement emphasizing the benefit of the brand.

Play with words: using typography of the words, symbols, or logos to represent the benefit or the USP of the brand.

Playing with the time: making the effects of time visible, put the brand in the spectrum of time either in the past, present, or future.

Provocation and shock tactics: challenging, inciting, and stimulating the audience.

Reframing: providing a key to creative thinking.

Repetition and accumulation: disrupting a regularly repeated pattern.

Spoofs and parodies: changing a well-known brand, cartoon, slogan, book hero, and so on, in order to conflict with original function and character, providing a new meaning.

Symbols and signs: using symbols, signs or combination of them to present brand benefit.

Take it literally: translate the words or phrases into a picture.

Telling stories: everyday stories around the brand following a dramatic style, such as adventure, love story, drama, documentary, news, or thriller.

Turn it right around: doing the opposite of what one would expect, using features of the brand, benefits, packaging in an opposite way.

Without words: telling stories in an effective way without using words, depicting USP without words or in a silent way, portraying the brand benefit in one picture (Priken, 2004).

Once the creative process is terminated the proposed campaign then passes to the evaluation stage.

Evaluating. Further research and testing methods to select the final campaign are then implemented. It is usually the client who is responsible for the final decision as to effectiveness.

Implementing. The plan is then executed.

Measuring. Results are measured and evaluated. Adjustments if required are implemented.

The role of creativity is clear as it has to result in an improvement in achieving organizational and more specifically marketing objectives. Effective advertising must contain a unique idea that captures consumers' attention, attracts their interest, and stimulates them to have a positive reaction toward the product. At the end of the day, creative advertising should facilitate the product or service to be discerned from any other creation including those of competitors.

Imminent Influencing Trends

Presently there appears to be a dilemma on the future of the actual creative processes and the advertising agency in general.

To analyze future trends, brand managers/ambassadors of global products were contacted. Surprisingly, the results appeared obvious as all managers/directors responded likewise. Permission was granted to cite one of the global distribution companies; Pernod Ricard, who deal in alcoholic beverages. Their Jameson Irish Whisky Ambassador, Graham Wood, said the following:

> The key role at the moment is the Media Shop because it is far too expensive to run an in-house creativity department. We are aware that there will be a big internet explosion. However, presently solid data/research is lacking.

The internet is most definitely the future. Presently we are seeing unpredicted phenomena. Sites such as *Facebook, YouTube, MySpace and Twitter* are rapidly increasing their membership. All these sites support the political, commercial, sportive, humanitarian, and personal messages.

Facebook will most probably emerge as the site for all age groups, nationalities, and genres (*see* SOCIAL INFLUENCE). The use of Facebook for sending political messages debuted with President Barack Obama. On his official page created prior election he had seven million 'friends' globally in less than a week. This is not including the 500 plus unofficial pages with over 30 million friends/fans. Future policies are

forwarded and all "friends" may comment, send ideas, and or question the policy. Commercial brands are slowly appearing officially and unofficially obtaining several thousand friends/fans in less than a day. Where sport is concerned the major European football clubs are presented again officially and unofficially. Humanitarian groups, for example; "Feed a child with just a click" has nearly five million members.

YouTube is a video site and previously catered to a younger audience. Since the debut of the share option, with a mere click members may send videos to their Facebook or MySpace pages and to email accounts. This has created a surge of new, all age members. The impact is enormous. SusanBoyle, a Britain's Got Talent contestant, did not win the competition but still became an overnight global star due to the site. Ten million people globally saw her video in a weekend. The Obama music video 'Yes We Can' had over 19 million hits in the same period. Advertisements and their spoofs are well positioned especially if they are humoristic.

MySpace members use their pages for storing photos, videos and playlists. This site is used mainly by a younger audience promoting their personal music and videos. It also permits "friends". Taken at random and contacted an unknown "Mr. Dialysis" a hip hop rapper has for example over 15,000 fans who left over 10,000 comments on his page. The site also features a contact icon which has permitted promoters to contact with gig offers. Mr. Dialysis states "My Space permitted me to get my music and lyrics out there".

Twitter, is the text "what am I doing right now" site and has a large following. However its appeal is diminishing amongst the older audience because of the "stalking" perception. The British Government found the site interesting and posted a generously paid job offer for a general public termed "twittercrat".

For marketers there will be the problem of keeping their page unique by somehow blocking the myriad of unofficial pages, or allowing these pages to propagate the message further. The danger lies in the unofficial who distorts the message or the embarrassment of a spoof that betters or negates the original.

Finally, it is evident that academic research is desperately needed, based not only on Internet possibilities, but which sites, what form of media, and how to "control" the message. Also an empiric study is also needed on Media Placement Specialists, an expanding business open to all due to increased Internet usage.

Bibliography

Amabile, M.T. (1983) *The Social Psychology of Creativity*, Springer-Verlag, New York.

Amabile, M.T. (1996) *Creativity in Context*, Westview Press, Boulder.

Ang, S.H. and Low, Y.M. (2000) Exploring the dimensions of ad creativity. *Psychology and Marketing*, **17**, 835–854.

Belch, E.G. and Belch, A.M. (2009) *Advertising and Promotion: An Integrated Marketing Communications Perspective*, 8th edn. McGraw-Hill, Irwin.

Burnett, L. (1968) Keep llistening to that wee, small voice, in *Readings in Advertising and Promotion Strategy* (eds M.A. Barban and C.H. Sandage), Richard D. Irwin, Homewood.

Csikszentmihalyi, M. (1999) Implications of a systems perspective for the study of creativity, in *Handbook of Creativity* 9ed. R. Sternberg), Cambridge University, New York.

El-murad, J. and West, C.D. (2003) Risk and creativity in advertising. *Journal of Marketing Management*, **19**, 657–673.

El-Murad, J. and West, C.D. (2004) The definition and measurement of creativity: what do we know? *Journal of Advertising Research*, **44**, 188–201.

Hirschman, E.C. (1989) Role-based models of advertising creation and production. *Journal of Advertising*, **18**, 42–53.

Kilgour, M. and Koslow, S. (2009) Why and how do creative thinking techniques work? Trading off originality and appropriateness to make more creative advertising. *Journal of the Academy Marketing Science*, **37**, 298–309.

Ogilvy, D. (1964) *Confessions of an Advertising Man*, Longman, London.

Ogilvy, D. (1985) *Ogilvy on Advertising*, Vintage Books, New York.

Oxford Advanced Learner's Dictionary, (2009) Oxford University Press, Oxford

Priken, M. (2004) *Creative Advertising; Ideas and Techniques from the World's Best Campaign's*, Thames & Hudson Ltd.

Reeves, R. (1961) *Reality in Advertising*, Alfred A. Knof. New York.

Smith, E.R. and Yang, X. (2004) Toward a general theory of creativity in advertising: examining the role of divergence. *Marketing Theory*, **4**, 31–58.

Smith, E.R., MacKenzie, B.S., Yang, X. *et al*. (2007) Modeling the determinants and effects of creativity in advertising. *Marketing Science*, **26**, 819–837.

Sternberg, J.R. and Lubart, I.T. (1996) Investing in creativity. *American Psychologist*, **51**, 677–688.

Till, D.B. and Baack, W.D. (2005) Recall and persuasion: does creative advertising matter? *Journal of Advertising*, **34**, 47–57.

West, C.D., Kover, J.A., and Caruana, A. (2008) Practitioner and customer views of advertising creativity – same concept, different meaning? *Journal of Advertising*, **37**, 35–45.

Young, W.J. (2003) *A Technique for Producing Ideas*, The McGraw Hill Companies, Inc., New York.

marketing functions on the Internet

Larry Chiagouris and Alexis Verniere

The commercialization of the Internet, which began in earnest in the mid-1990s, has generated substantial opportunity. Prior to 1996, however, there was no significant body of research that could be relied upon for direction on how to take advantage of the promise of the Internet. One of the most extensive research reviews found that through 2003, there still remained much to be learned and shared about how to best conduct business on the Internet (Cho and Khang, 2006).

Much has changed since 1996. Marketing functions that are conducted in the traditional brick-and-mortar world also can be found operating within the Internet. Managers have now translated marketing functions to work in an on-line environment, including functions such as market research, product development, branding, communications, pricing, and retailing. They have done so while having to develop realistic business models and addressing consumer privacy issues. The following examines each of these areas.

Market Research

It has been estimated that the majority of quantitative data collected in the United States is now collected on-line. Qualitative research is also growing by connecting people with a moderator on-line in order to explore marketing and product-related issues.

The documentation of using the Internet to conduct market research can be traced to the very early period of the commercialization of the Internet (see Berry *et al.*, 1994). Internet market research has had a tremendous impact on how marketing decisions are made. Market research on-line can be conducted with greater speed, lower costs, and improved accuracy as compared to off-line methods (Lockett and Blackman, 2004). Internet market research allows consumers to examine products, concepts, and test advertising from the comfort of their own home and provide their reactions as to what they like and dislike about products and tactics. Importantly, because of the declining rates of cooperation in off-line research, on-line research is becoming the most common form of market research.

Key to the growth of the use of market research over the Internet have been research studies that were launched to compare the results of on-line research methods with the more established off-line methods. While there can be some important differences, credible research has found that the results produced via the web are comparable to what would be expected in using more traditional off-line methods (Kaplowitz, Hadlock, and Levine, 2005).

One way that on-line research is being used is through feedback obtained via on-line communities. These communities are being formed through social networking sites such as MySpace and Facebook. Research with on-line communities allows companies to become better acquainted with customers and identify and leverage new trends in a timelier fashion (Casteleyn, Mottart and Rutten, 2009).

Similar to other research methods, market research conducted over the Internet has several issues and challenges. The issues and challenges include too much data to decipher, low response rates, and the delay in implementing new methods of market research.

A growing body of research and literature has emerged as to what to do and what to avoid in order for on-line research to be conducted effectively. High on the list of suggestions is the need to keep the survey short, and assure that it has been properly pretested and conducted among a representative sample of the external audience that it is intended to represent.

Product Development

The Internet has greatly enhanced the participation by consumers in the product development process. Through user-friendly development tools and the expansion of broadband service, customer participation in product development has changed the process of product innovation.

The benefits to marketers include lower product development costs, faster development times, and increased customer relevancy of new products. There is also some evidence that the Internet increases demand for products early in the product launch phase (Prince and

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Simon, 2009). This is due to consumers having improved access to information in addition to the convenience of on-line shopping.

There are many examples of this form of product innovation. Companies have been able to develop and distribute software-based design tools or "tool kits" that allow lead users to work with the companies to make new products or refine existing products and services. For example, more than 80% of computer game companies have on-line user communities that are used to develop new product opportunities (Jeppesen and Molin, 2003).

An extension of the Internet in the product development process is the concept of virtual reality. Here, consumers are able to project themselves into three-dimensional spaces. They can interact with products (through an avatar that represents their persona) under the supervision of the product development team.

This use of virtual reality is becoming recognized as holding great potential to impact product development programs in the future (see Ottosson and Holmdahl, 2007; Mackenzie, Buckby, and Irvine, 2009). Some companies, however, are concerned that they may lose control of the product development process to their user base. Therefore, these companies are seeking to reign in user-driven development efforts over the Internet (see Braun and Herstatt, 2008).

Branding

One of the first articles to address branding on the Internet (see Chiagouris and Wansley, 2000) addressed the challenge of how a brand can transition prospects to customers and build relationships. The Internet has major advantages over mass media. Mass media cannot communicate with individual prospects in a customized way based on the different stages of the relationship-building process.

Rather than simply interrupting a television show with a commercial message or barging into their lives with unannounced phone calls or letters, innovative marketers first try to get individuals to voluntarily enroll in the relationship-building process. A volunteer's experience will almost always be more likely to result in the embrace of a brand than any forced viewing experience. Interactive technology enables marketers to inexpensively attract consumers into one-to-one relationships fueled by two-way "conversations" played out via mouse clicks on a computer.

There are many design elements that can accomplish the objectives of moving a prospect along the brand relationship-building continuum. These techniques range from brief registration forms that begin the relationship-building process to games that intrigue the user on a category-relevant subject. There are also programs designed to identify if the user has visited the site before and greets/treats the user in a manner that reflects this knowledge.

Amazon.com provides an excellent example of how to build a relationship on the Internet. Once a customer buys a book from it, Amazon.com begins collecting information and making educated guesses about what might appeal to this person. Then it makes book selections and gently guides the individual toward other choices. It does so by creating communities of special interest groups (e.g., art, cooking, or sports) and offering appropriate links throughout the site. Today when a customer logs on to their website, Amazon.com provides a message of "personal recommendations" which is based on selection criteria.

Traditional marketers like Procter & Gamble (P&G) have also engaged in Internet branding. P&G also demonstrated creativity in their Internet branding campaigns. One of their most successful interactive campaigns was the Scope e-mail kiss. The campaign generated an average click-through rate of 20%, with an even higher conversion rate of pass-along e-mails.

While Internet branding shares much with traditional branding methods, there are some important distinguishing characteristics. Internet branding consists of fluid strategies regarding new products and campaigns rather than finite starting and ending points. Internet branding does not have the tangible collateral that off-line branding requires. Importantly, Internet branding addresses the individual interests of the customer rather than treating all customers alike.

Although the theoretical foundation of branding on the Internet is still in an emerging stage, scholars have identified key issues and

suggestions for effective branding. For example, providing the appropriate cues and symbols for a brand on-line can contribute to a more effective brand presence. In particular, building brand alliances with respected and established brands can contribute to a more effective brand presence (Delgado-Ballester and Herrnandez-Espallardo, 2008). This is particularly the case if the brand alliance is promoted on a brand website.

In a similar fashion, use of trustmarks, such as BBBOnline or TRUSTe can be a more effective tool in boosting brand imagery than ratings from magazine articles or paid advertising support. These third-party graphic images placed on a website can decrease the anxiety that shoppers have about purchasing something on-line. These marks add to the quality of the brand imagery of an on-line brand.

The measurement of on-line brand equity is also becoming increasingly important. Scholars have found that there are a number tactics that will enhance on-line brand equity. These tactics range from the creation of on-line newsletters and magazines (Müller *et al.*, 2008) through to on-line programs that telegraph personalization (DaSilva and Faridah, 2008).

Marketing Communications

The Internet, as an advertising class, provides some of the most varied forms of marketing communications. The type and nature of forms of marketing communications is highly varied and constantly evolving. According to a 2009 report from investment bank Morgan Stanley, marketers now consider this form of communications more effective than television or newspaper advertising. Internet advertising is expected to receive 15% of all global advertising spending in 2010 (Group, 2009).

Early in its development, the absence of standardized creative units presented many challenges to the marketing and advertising industry. As a result, the Interactive Advertising Bureau (then called the *Internet Advertising Bureau*) worked with on-line publishers, advertisers, agencies, and media companies to create a set of voluntary guidelines. These voluntary guidelines are now in use in creating, planning, buying, and selling of interactive marketing and advertising.

The standards address a wide variety of advertising formats, including rectangles, pop-ups, banners, buttons, and skyscrapers. The standards specify the sizes in terms of pixels as well as file weight and maximum time for the loading of rich media (ads that use advanced technology such as streaming video or programs that interact with the user).

Marketers seeking to place ads on the Internet can do so through a variety of ways. The most popular ways are to work with an ad network or directly with an on-line publisher. Recent research indicates that if carefully chosen, working with an on-line publisher directly is likely to lead to better results. On-line publishers are a critical component of marketing-related considerations and marketers often partner with on-line publishers to achieve their objectives.

The principal functions of an on-line publisher that operates a website on the Internet consists of creating content, producing it in a manner that it can be uploaded to its website for viewing by the people who visit its website, and then generating traffic in the form of visitors to its website. Advertising revenue is the most common form of revenue for on-line publishers. Advertising revenue is generated when people who visit the site are exposed to the ads on the site from advertisers. Usually, the advertisers will compensate the operator of a website that displays their ads based on the number of people who have been exposed to the advertisers' ads.

The conditions under which banner advertising has been found to be effective vary. For example, brand recall is higher for a banner ad when the ad is relevant to the content found on the website in which it appears (Yaveroglu and Donthu, 2008). The Internet is evolving into a mature medium, however, and it has exhibited an effect referred to as *banner blindness* (Cho and Khang, 2004). This effect is a condition in which consumers ignore banner ads found on website. As a result of some erosion in the effectiveness of banner ads, marketers are increasingly using search engine marketing (discussed below) and e-mail marketing in their marketing communications plans.

In terms of e-mail marketing, 50% of consumers indicate that they are more likely to purchase products from companies that send them e-mail (Epsilon Branding Survey, 2009).

Two-thirds of consumers have indicated that they have purchased something because of a message that they have received within an e-mail (ExactTarget, 2008). E-mail marketing is now more than a $1 billion industry. As such, a substantial amount of attention is being given to e-mail marketing programs by marketers. Effective e-mails often contain information of interest to consumers, sometimes coming in the form of a newsletter. The most effective forms of e-mail marketing are the ones that utilize an opt-in process to build the e-mail list.

Single opt-in occurs when a consumer clicks on a link or a button and indicates a willingness to receive an e-mail from a particular company. Double opt-in is considered to be even more effective. This occurs after a consumer clicks on a link or button. That process is, in turn, followed by the consumer receiving an e-mail in which they then have to click on a link within the e-mail to increase the certainty that the consumer wishes to receive e-mails from a particular company. All must abide by spam regulations, which require the need to allow consumers to unsubscribe to an e-mail solicitor.

Advertising on the Internet has evolved to take on many new forms and on devices other than computers. In terms of content, social networks, blogs, and other techniques that leverage user-generated content are becoming more popular. There are some lingering questions, however, concerning whether these newer forms of content will have sustainability. Revenue for MySpace declined 16% in the first quarter of 2009. Bloggers who are at times considered a part of the journalism sector have come under closer scrutiny for bending principles of traditionally accepted journalistic ethics.

In terms of devices, the mobile phone has captured the attention of marketers. Influential articles that address mobile phones as a marketing tool began to emerge in 2005 (see Sultan and Rohm, 2005). This form of advertising is now adapting Bluetooth technology to provide sponsored alerts to prospective customers.

Mobile phones are also supporting short message service (SMS) capabilities that facilitate advertiser messages. Sponsored entertainment over smart phones with large screens can be expected to become a popular application in the future. The key challenge for some advertisers, however, is to create websites that employ the wireless application protocol (WAP) so that the appearance of the website on the mobile phone is easy to comprehend and the features easily used. The industry will also need to be sensitive to some consumers finding advertiser-supported messages intrusive on their cell phones.

Additional challenges are expected to confront advertisers in the future. One challenge is that consumers using the Internet may not always be attentive to it. Nielsen reports that more than half of all consumers are doing something else when on the Internet. An environment in which consumers are multitasking when interacting with the Internet is currently a major challenge to advertisers in seeking to get the most return for their investments in the Internet.

Pricing

The Internet has made it increasingly easier for consumers to know at both wholesale and retail levels the prices that are being paid for goods and services. While all traditional brick-and-mortar pricing strategies apply to the Internet, there are some approaches to pricing that are either specific to or greatly enhanced in an Internet business.

Dynamic pricing is a term for a pricing strategy in which the seller changes prices either over time or across products or segments of consumers (see Kannan and Kopalle, 2001; Jallat and Ancarani, 2008; Kumar and Sethi, 2009). The Internet has made it easier for both parties to a transaction to adjust willingness to pay (buyer) and to adjust the price of the product or service (seller).

Dynamic pricing can take a variety of forms with two of the most popular being the updating of prices by the seller (e.g., airlines) or the arrangement of an auction in which potential buyers bid against each other to acquire a product within a specific time frame. With regard to auctions, the supplier can usually expect to pay a commission to the entity arranging for the auction and that commission is usually factored into its cost structure and ultimate price point.

In terms of auction-site pricing, there is research that suggests that consumers in certain parts of the country are more willing to pay

higher prices than those in other parts of the country (Bland, Black, and Lawrimore, 2007). In the same research, there is an indication that women might be willing to pay more than men for a product in an auction environment. Willingness to pay may also decline when similar items priced at different levels are listed in an auction environment or when the prospective bidder has an extensive bidding history.

Variable pricing is yet another form of pricing made possible by the Internet. In this case, the consumer may not be willing to pay a listed price. Here, the price can be lowered if the consumer is willing to accept certain limitations. In a popular example at Priceline.com, the consumer may have to travel on a flight or at a time of day determined by the supplier. In another example, progressive insurance, the consumer names his or her price and some key attributes but then accepts the remaining terms and conditions of the insurance policy.

Another pricing strategy is the cost-of-acquisition pricing strategy. It is based on a determination of the cost to acquire a visitor to a website and then the probability that the visitor will purchase something. On the basis of this cost, which is based primarily on the cost to drive traffic to a website, the business applies its margin in order to reach profit objectives.

Digital products and services often have low or zero marginal costs. For these, a value-added pricing strategy will often be used. The price is not driven by costs but by how much value the seller believes that the buyer will receive. This "value" is often determined through the use of marketing research studies.

On-Line Retailing and Customer Service

According to the US Census Bureau, e-commerce sales in the United States was more than $125 billion in 2009. Analysts at Goldman Sachs expect more than 15% growth in e-commerce sales in 2010. The majority of consumers have now become on-line buyers. Internet retailing has become mainstream.

The growth in on-line sales is particularly robust, given that the shopping cart abandonment rate was 48% in 2005 (eMarketer, 2006) and some estimates put the rate of abandoning an on-line shopping cart as high as 75%. The reasons that consumers provide for not completing an on-line transaction are quite varied and on-line retailers are continuing to work toward overcoming prospect concerns. The opportunities and concerns incorporate the design of online stores, changing attitudes and behavior and providing strong customer service and relationship-building programs online.

In terms of on-line store design, there are a wide variety of considerations related to retail store design in the brick-and-mortar sector that have on-line equivalents. Similar to the off-line environment, the on-line retail environment must have an atmosphere that can make a difference to prospective shoppers (Eroglu, Machleit, and Davis, 2003). The on-line retailer can adjust, manage, and manipulate a variety of cues that contribute to the shopping experience. Many visual and auditory stimuli can be addressed to enhance the on-line shopping experience.

The consumer's attitude toward a website is important in his or her intention to purchase from it or come back to it. Factors affecting attitude toward a retail website include ease of use, product information, entertainment, trust, and currency – having everything on it up to date (Elliott and Speck, 2005). Importantly, not all Internet visitors react the same way to a website. For example, in the Elliot and Speck study, low-involvement shoppers who rated a website entertaining were also likely to rate it favorably. High-involvement shoppers were more interested in how current a website was, and rated it favorably if everything was up to date.

Attitude toward a website also differs depending on the shopper's level of experience. Inexperienced shoppers are likely to favorably rate websites that they perceived as easy to use. Conversely, product information on a website appeared to sway experienced shoppers more. They favorably rated websites that fit their information needs.

Creating trust is an important consideration in on-line retailing. If a consumer does not trust the offerings on a website, then he or she will not purchase anything from it. Dimensions of website design that contribute to consumers' perceptions of trust include visual website design, content design, and social-cue design (Wang and Emurian, 2005). The last item

encompasses all of the potential for consumer interaction with a website, as well as the use of multimedia.

Consumers with different levels of experience can be expected to react to on-line retailing offers in a way that mirrors their on-line shopping skills. Experienced Internet shoppers rely more on the efficiency of the shopping experience, 24-h availability and the desire to avoid waiting lines/queues (Karayanni, 2003). In another work, it was found that more experienced shoppers found merchandizing tactics more impactful at building on-line trust, while inexperienced shoppers rely more on fulfillment, promotion, and website design (Jin, Park, and Kim, 2008).

An important component of any on-line selling environment is the customer service that is provided for visitors. In a recent survey by Harris Interactive and Tealeaf Technology, it has been reported that 57% of on-line shoppers would be less likely to buy from a brand off-line if they experienced problems on-line (McEleny, 2008). In effect, providing effective on-line customer service not only impacts Internet sales but it also has the potential to influence off-line transactions.

Underscoring the linkage between on-line and off-line businesses, it was also reported that 89% of consumers believe that on-line customer service should be the same level or better than shopping in a store. Importantly, 49% of consumers who experience problems with on-line transactions indicate that they would abandon or switch the transaction to another site – up from 37% in a similar survey conducted in 2007.

One emerging field to address customer service issues and concerns is electronic commerce customer relationship management programs. These programs are emerging as an important part of managing customer service programs over the Internet (Ratnasingam, 2008). The programs incorporate the use of applications to manage customers and customer interactions on the web.

Properly used, customer relationship management programs allow a company to reduce the costs of communicating with customers, improve customer interaction with the firm, and lead to service improvements. Marketers are recognizing its importance and are investing in related software. For example, estimates provided by the Gartner Group indicate that spending on on-line community social networking technologies software neared $400 million in 2008.

Discontent with service recovery by customers has been observed to be increasing (Holloway and Beatty, 2003). Responding to service problems in an on-line environment, however, has been found to be uneven across companies. There apparently are differences between how large companies and small companies address on-line complaints. Larger companies appear to provide more channels for their customers to register complaints compared to smaller companies (Neale, Murphy, and Scharl, 2006).

Recognizing that there are differences between on-line and off-line environments, measures of service quality are beginning to emerge that have been developed specifically for the on-line environment. These scales incorporate the dimensions of efficiency, fulfillment, system availability, privacy, responsiveness, compensation, and contact (see Parasuraman, Zeithaml, and Malhotra, 2004). One scale under development is the "eTail" quality or "eTailQ" (Wolfinbarger and Gilly, 2003). Website design, fulfillment/reliability, privacy/security, and customer service have been identified as predictive of eTailQ.

Internet Privacy

On-line businesses often utilize extensive data-gathering methods in order to target the right customers. These methods include data mining and conducting surveys. These methods also include a technique referred to as *behavioral targeting*. Behavioral targeting uses information collected during the browsing session by a visitor to a web page. This is accomplished by a website page placing a cookie (small file) on the hard drive of a consumer's computer. This cookie is the basis for sending information back to the company or companies involved in tracking consumer behavior.

Companies can use information generated through behavioral targeting to identify what pages are visited by consumers. Marketers expect

that this will aid them in delivering their on-line advertisements to the users who are most likely to be interested in the advertisements. Behavioral targeting is often used in conjunction with other forms of targeting to include information based on factors like past purchases, geography, and demographics. The assumption is that targeted ads will generate higher levels of purchase interest in consumers.

Personalization can be a very effective method to market to specific customers. Even though the results from personalization can be beneficial, the act of personalizing marketing efforts over the Internet may be seen as intrusive or an invasion of privacy. Over the last decade, privacy concerns of on-line users have been on the rise (Yang and Wang, 2009).

Behavioral targeting, in particular, is viewed by some consumers and consumer groups as an invasion of privacy. As a result of consumer push back, Internet privacy laws that restrict the use of behavioral targeting may be enacted in the future. Lawmakers have not reached a consensus as to what are the key threats related to Internet privacy. There is also a question as to how far to extend legislation and specifically, whether or not to limit legislation exclusively to e-commerce.

Privacy concerns are not equally shared by consumers. For example, gender can be a moderator variable in terms of the level of consumer privacy concerns. Recent research found that girls tend to exhibit greater privacy concerns than boys (Seounmi and Hall, 2008). In addition, girls and boys differ with regard to how much and what kind of information they share in the process of registering with a website.

Business Models and Traffic Generation

Websites designed for selling products and services or informing and entertaining have the ultimate objective of providing a return on investment. Many early Internet businesses did not have business models that produced profit. As such, discussions concerning business models in connection with the Internet have been extensive. While there are many variations and combinations of business models, a few have emerged to become more recognized as effective ways to conduct business on the Internet.

One form of revenue generation is a business model based on e-commerce. In this form of revenue generation, an Internet organization sells products or services and takes some portion of the sale price as revenue. This model often requires substantial negotiation concerning prices paid and commissions earned.

Another form of revenue generation is based on what is called a *subscription business model*. Using this form of business model, an Internet organization charges visitors a subscription fee in order to access information or services provided by it. Well-established newspapers in particular, have been seeking to transition to this form of revenue generation. Financial services organizations that provide stock-picking tips also are seeking to adopt this business model.

Pay per click is another form of revenue generation in which search engines play an important role. Internet organizations receive revenue from advertisers or the agents of advertisers (search engines). In return, the Internet organizations allow advertisers' ads to be placed on their respective websites so that visitors to their websites will be exposed to the advertisers' ads.

Pay per click creates revenue if a visitor clicks on an ad that he or she is exposed to when visiting a website. The operator of the website receives a fee for each visitor who clicks on the advertiser's ad. Text ads on search engines are often the form of advertising used and "clicked on". Google, Yahoo!, and other search engines often use this form of advertising revenue generation. Operators of websites also allow Google to place ads on their respective sites. If a visitor clicks on these ads, Google receives a fee from the advertisers of the sponsored ads. Google then pays each of its participating sites a portion of that fee for allowing the placement of Google provided ads on their sites.

Pay per click is often used to generate traffic to a website. Fee for traffic is the money that one business on the Internet pays another business (or person) on the Internet in return for the second business sending people to visit to visit the website of the first business. The reason why the first business described above is willing to pay a fee for this traffic is that it expects to generate more business directly related to the

number of people who visit its website. It is important to note that the amount of the fee is directly related to the amount of the traffic. In other words, fee for traffic is customarily priced in such a way that the total fee charged is based on the total number of visitors and amount of traffic sent by one entity to another entity.

Businesses often will place links on their sites to other sites. They do so in anticipation that people who visit their sites will click on the links and then be taken to a different site. If that happens, the second site pays the first site a fee for the traffic. The name often given to this relationship between the two sites is an *affiliation* and it is often formalized in what is called an *affiliation agreement*.

In considering tools to generate traffic generation, pay per click is quite popular and is an important part of the area of practice referred to as *search engine marketing*. Another complementary tool is search engine optimization. This takes the form of an operator of a website modifying the words and images that appear on its site so that search engines will increase the likelihood that the website will be listed in its results that appear to a user who searches for something on the search engine. These results appear in the section of results that are not sponsored or the section sometimes referred to as *organic results*. The challenge here for the website operator is to constantly manage and modify its search engine optimization efforts in order to remain popular with search engines.

Bibliography

Berry, J., Verity, J., Kerwin, K., and De George, G. (1994) Database marketing. Business Week, 56–62.

Bland, E., Black, G., and Lawrimore, K. (2007) Risk-reducing and risk-enhancing factors impacting online auction outcomes: empirical evidence from ebay auctions. *Journal of Electronic Commerce Research*, **8** (4), 236–243.

Braun, V. and Herstatt, C. (2008) Barriers to user innovation: moving towards a paradigm of 'licence to innovate'? *International Journal of Technology Policy and Management*, **7**, 292–303.

Byoungho, J. and Park, J.Y. (2006) The moderating effect of online purchase experience on the evaluation of online store attributes and the subsequent impact on market response outcomes. *Advances in Consumer Research*, **22**, 203–211.

Casteleyn, J., Mottart, A., and Rutten, K. (2009) How to use Facebook in your market research. *International Journal of Market Research*, 2009, **51**, 439–447.

Chiagouris, L.G. and Wansley, B. (2000) Branding on the Internet. *Marketing Management*, **9**, 34–38.

Cho, C. and Khang, H.J. (2004) Why do people avoid advertising on the internet? *Journal of Advertising*, **33**, 89–97.

Cho, C. and Khang, H. (2006) The state of internet-related research in communications, marketing, and advertising: 1994-2003. *Journal of Advertising*, **35**, 143–163.

DaSilva, R.V. and Faridah, S. (2008) Online brand attributes and online corporate brand images. *European Journal of Marketing*, **42**, 1039–1058.

Delgado-Ballester, E. and Hernández-Espallardo, M. (2008) Building online brands through brand alliances in Internet. *European Journal of Marketing*, **42**, 954–976.

Elliott, M.T. and Speck, P.S. (2005) Factors that affect attitude toward a retail web site. *Journal of Marketing Theory and Practice*, **13**, 40–51.

eMarketer. (2006) US retail e-commerce shopping cart abandonment rates, 2001-2005. Retrieved from: http://www3.emarketer.com/Results.aspx?N=1068&No=1200

Epsilon Branding Survey (2009) Retrieved from: http://www.emailstatcenter.com/Branding.html

Eroglu, S.A., Machleit, K.A., and Davis, L.M. (2003) Empirical testing of a model of online store atmospherics and shopper responses. *Psychology and Marketing*, **20**, 139–150.

ExactTarget (2008) Channel Preference Survey, Retrieved from: http://email.exacttarget.com/Resources/Whitepaper/ 2008_Channel_Preference_Survey.html

Group, M. (2009) Study says Internet advertising spending will reach 15% of total in 2010. Retrieved from: http://www.webwire.com/View-PressRel.asp?aId=104187

Holloway, B.B. and Beatty, S.E. (2003) Service failure in online retailing: a recovery opportunity. *Journal of Service Research*, **6**, 92–105.

Jallat, F. and Ancarani, F. (2008) Yield management, dynamic pricing and CRM in telecommunications. *The Journal of Services Marketing*, **22**, 465–478.

Jeppesen, L.B. and Molin, M.J. (2003) Consumers as co-developers: learning and innovation outside the firm. *Technology Analysis and Strategic Management*, **15**, 363–384.

Jin, B., Park, J.Y., and Kim, J. (2008) Cross-cultural examination of the relationships among firm reputation, e-satisfaction, e-trust, and e-loyalty. *International Marketing Review*, **25**, 324–337.

Kannan, P.K. and Kopalle, P.K. (2001) Dynamic pricing on the internet: importance and implications for consumer behavior. *International Journal of Electronic Commerce*, **5**, 63–83.

Kaplowitz, M.D., Hadlock, T.D., and Levine, R. (2005) A comparison of web and mail survey response rates. *Public Opinion Quarterly*, **68**, 94–101.

Karayanni, D.A. (2003) Web-shoppers and non-shoppers: compatibility, relative advantage and demographics. *European Business Review*, **15**, 141.

Kumar, S. and Sethi, S. (2009) Dynamic pricing and advertising for web content providers. *European Journal of Operational Research*, **197**, 924–944.

Lockett, A. and Blackman, I. (2004) Conducting marketing research using the Internet: the case of Xenon Laboratories. *Journal of Business and Industrial Marketing*, **19**, 178–187.

Mackenzie, K., Buckby, S., and Irvine, H. (2009) A framework for evaluating business lead users' virtual reality innovations in Second Life. *Electronic Commerce Research*, **9**, 183–202.

McEleny, C. (2008) Poor ecommerce puts customers off physical stores too. *New Media Age*, 10.

Müller, B., Florès, L., Agrebi, M., and Chandon, J.L. (2008) The branding impact of brand websites: do newsletters and consumer magazines have a moderating role?. *Journal of Advertising Research*, **48**, 465–472.

Neale, L., Murphy, J., and Scharl, A. (2006) Comparing the diffusion of online service recovery in small and large organizations. *Journal of Marketing Communications*, **12**, 165–181.

Ottosson, S. and Holmdahl, L. (2007) Web-based virtual reality. *Journal of Engineering Design*, **18**, 103–111.

Parasuraman, A., Zeithaml, V.A., and Malhotra, A. (2004) E-S-Qual: a multiple-item scale for assessing electronic service quality. *Marketing Science Institute Working Paper*, 04–112.

Prince, J.T. and Simon, D.H. (2009) Has the Internet accelerated the diffusion of new products?. *Research Policy*, **38**, 1269–1277.

Ratnasingam, P. (2008) The impact of e-commerce customer relationship management in business-to-consumer e-commerce. *Journal of Electronic Commerce in Organizations*, **6** (4), 30–46.

Seounmi, Y. and Hall, K. (2008) Gender and online privacy among teens: risk perception, privacy concerns, and protection behaviors. *CyberPsychology and Behavior*, **11**, 763–765.

Sultan, F. and Rohm, A. (2005) The coming era of 'brand in the hand' marketing. *MIT Sloan Management Review*, **47**, 83–90.

Wang, Y.D. and Emurian, H.H. (2005) Trust in e-commerce: consideration of interface design factors. *Journal of Electronic Commerce in Organizations*, **3**, 42–60.

Wolfinbarger, M. and Gilly, M.C. (2003) eTailQ: dimensionalizing, measuring and predicting eTail quality. *Journal of Retailing*, **79**, 183–198.

Yang, S. and Wang, K. (2009) The influence of information sensitivity compensation on privacy concern and behavioral intention. *Database for Advances in Information Systems*, **40**, 38–51.

Yaveroglu, I. and Donthu, N. (2008) Advertising repetition and placement issues in on-line environments. *Journal of Advertising*, **37**, 31–43.

comparative advertising

Larry Chiagouris

Pepsi versus Coke, Mac versus PC's, Dunkin' Donuts versus Starbucks, or Avis versus Hertz are examples of comparative advertising campaigns. Burger wars, Pepsi Challenges, and other brand comparisons are highly visible and have captured the attention of the media and consumers. This is an examination of the development of comparative advertising and conditions under which it can have more of an impact on marketing results.

Comparative advertising occurs when one brand provides a mention of another brand or brands in its advertising or in its related presentations and collateral materials. The comparison can be implied or can be specific to a particular brand or a set of brands.

The origin of comparative advertising can be traced to hundreds of years ago in England (Swayne and Stevenson, 1987). More recently, in 1930, Sears ran ads comparing its brand of tires to eight other national brands. Shortly thereafter, Plymouth ran advertising claiming that its cars were superior to Chevrolet and Ford.

It was not until 1971, however, that the number of comparative advertising campaigns in use increased substantially. This was due to a memo sent to the ABC and CBS Television Networks by Robert Pitofsky, the then Director of Consumer Protection at the Federal Trade Commission. In that memo, he indicated that consumers would benefit from explicit product and brand comparisons rather than indirect comparisons, which he believed might be misleading and deceptive. He encouraged the networks to aggressively seek out and accept comparative advertising campaigns.

Comparative advertising campaigns have been the source of substantial industry debates and scholarly research as to how and in what manner comparative advertising impacts marketing results. The reasons why comparative advertising might contribute to greater results have to do with speculation that comparative advertising receives more attention from consumers. First, comparative advertising might increase personal relevance to more consumers because the advertising names more than one brand (Wilkie and Farris, 1975). In addition, comparative advertising tends to contain more information than noncomparative advertising.

Advertisers often have to make a decision as to when to use a comparative advertising format as opposed to a noncomparative advertising format. Two of the most comprehensive reviews of comparative advertising resulted in conclusions that comparative advertising is not always more effective than noncomparative advertising. The relative effectiveness of the two formats depends on a variety of variables or conditions or types of comparative advertising formats (Barry, 1993; Grewal *et al.*, 1997).

One consideration is the degree that the key brand attribute used as a basis of comparison is typical to a product category. When the attribute is typical, a noncomparative ad format is more persuasive. This is not the case, however, when the attribute is atypical (Pillai and Goldsmith, 2008).

Information-processing styles of the target audience is also a potential consideration in the effectiveness of a comparative advertising campaign (Thompson and Hamilton, 2006). When consumers use an analytical processing (information processing that is data driven and more focused on verbal retrieval and encoding), comparative ads are likely to be more effective than noncomparative ads. For consumers who tend to use an imagery-based processing style (information processing that is based more on nonverbal sensory representations), noncomparative ads are likely to be more effective.

Individual differences linked to information processing styles also factors into the reaction to comparative versus noncomparative advertising formats. For example, gender of the consumer may play a role in the reaction to comparative versus noncomparative advertising (Chang, 2007). Men experience greater brand evaluation involvement than women and, as such, comparative appeals lead men to experience higher levels of brand evaluations and purchase intentions. Women, on the other hand, experience a greater sense of manipulative intent generated by comparative appeals and, as such, women may record more negative ad and brand evaluations and reduced purchase intentions when exposed to comparative advertising.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

In examining the potential effects of a comparative advertising campaign, the nature of the type of comparison may be important. One way of classifying a campaign is whether it is a direct comparison (e.g., "Honda is better than BMW"; "Miller is better than Bud") or an indirect comparison (e.g., "Motrin is better than the leading brand"; "Nobody delivers faster than FedEx"). Research has demonstrated that the nature of the comparison may produce different results moderated by variables such as category size and the perceived variability of the category (Miniard *et al*., 2006; Xiaojing *et al*., 2007).

Given the wide variety of variables, there is not one single number that represents an overall assessment as to the effectiveness of comparative advertising campaigns. There are estimates that properly executed comparative advertising campaigns can boost sales by as much as 15% (MacArthur and Cuneo, 2007). Belief in the format has resulted in many marketers adapting comparative advertising claims and it has also generated substantial litigation and disputes concerning competitive claims.

It has been recently reported that disputed ad claims related to comparative advertising have risen 40% in 2008 compared to 2007 (Wasserman, Hein, and Wong, 2009). There are several sets of rules and regulations governing the use of comparative advertising claims that may result in claims by one company against another in the use of a comparative advertising campaign. Companies often use these standards as a basis for bringing complaints against competitors.

Standards and guidelines relevant to comparative advertising include those from the American Association of Advertising Agencies, the United States Federal Trade Commission, and the National Advertising Division of the Better Business Bureau. Media vehicles, including the television networks, usually have their own guidelines.

In addition, governmental bodies at all levels are likely to have rules governing the practice of the use of comparative advertising. Section 43(a)(1) of the Lanham Act also provides a description of the conditions that a plaintiff must establish in seeking relief under the act. This act is often used as the basis for one party suing another party with the complaint noted as a violation of the Lanham Act.

Bibliography

Barry, T.E. (1993) Comparative advertising: What have we learned in two decades? *Journal of Advertising Research*, **33**, 19–29.

Chang, C. (2007) The relative effectiveness of comparative and noncomparative advertising: evidence for gender differences in information-processing strategies. *Journal of Advertising*, **36**, 21–36.

Grewal, D., Kavanoor, S., Fern, E.F. *et al*., (1997) Comparative versus noncomparative advertising: a meta-analysis. *Journal of Marketing*, **61**, 1–15.

MacArthur, K. and Cuneo, A.Z. (2007) Why big brands are getting into the ring. *Advertising Age*, **78**, 6.

Miniard, P.W., Barone, M.J., Rose, R.L., and Manning, K.C. (2006) A further assessment of indirect comparative advertising claims of superiority over all competitors. *Journal of Advertising*, **35**, 53–64.

Pillai, K.G. and Goldsmith, R.E. (2008) How brand attribute typicality and consumer commitment moderate the influence of comparative advertising. *Journal of Business Research*, **61**, 933–941.

Swayne, L.E. and Stevenson, T.H. (1987) Comparative advertising in horizontal business publications. *Industrial Marketing Management*, **16**, 71–77.

Thompson, D.V. and Hamilton, R.W. (2006) The effects of information processing mode on consumers' responses to comparative advertising. *Journal of Consumer Research*, **32** (4), 530–540.

Wasserman, T., Hein, K., and Wong, E. (2009) For some, 'Taste Test' ads leaving a bad taste. *Brandweek*, **50**, 8.

Wilkie, W.L. and Farris, P.W. (1975) Comparison advertising – problems and potential. *Journal of Marketing*, **39**, 4–11.

Xiaojing, Y., Jain, S.P., Charles Lindsey, C., and Kardes, F. (2007) Perceived variability, category size, and the relative effectiveness of 'Leading Brand' versus 'Best in Class' comparative advertising claims. *Advances in Consumer Research*, **34**, 209.

a review of ethnic identity in advertising

Jeremy J. Sierra, Michael R. Hyman, and Robert S. Heiser

What is Ethnic Identity and How Does it Relate to Advertising?

Ethnic identity entails self-identification as an ethnic group member, a sense of belonging to this group, and favorable attitudes toward this group (Phinney, 1992); it can be used as a way to overcome ad clutter and influence viewer responses to the traditional hierarchy-of-effects model in advertising. As ad impressions continue to mount, advertisers' difficulty in reaching their targeted consumers increases. One approach that may help overcome this difficulty entails the inclusion of ethnically resonant cues in ads, which may lead viewers to ethnically identify with such stimuli (Green, 1999); such methods have been shown to generate favorable viewer responses toward the ad and advertised brand (Sierra, Hyman, and Torres, 2009). Also, ads that depict ethnic characters, images, and values may lead ethnically resonant consumers to identify with the featured brand (Koslow, Shamdasani, and Touchstone, 1994).

Measuring Ethnic Identity in Advertising Studies

Ethnic identity is an enduring, underlying sense of connection to a social group (Tajfel, 1978). Strength of ethnic identity (i.e., strength of identity with one's ethnic origin (Deshpandé, Hoyer, and Donthu, 1986) or enduring association between one's ethnicity and sense of self (Forehand and Deshpandé, 2001)) can affect consumers' responses to marketing activities, shopping orientations, and product/service choices (Green, 1999).

Adolescents are more likely to believe and identify with ads laced with ethnically resonant cues (Appiah, 2001b). Relative to Blacks with weaker ethnic identities, Blacks with stronger ethnic identities see themselves as more similar to, and identify more strongly with, Black characters in ads (Appiah, 2001b; Whittler, 1989). Also, strong Hispanic identifiers prefer media that use Spanish verbiage (Deshpandé, Hoyer, and Donthu, 1986), which suggests that strength of ethnic identity relates positively with tendencies to notice and identify with ethnically resonant ads.

Although careful in-depth questioning could reveal respondents' ethnic identity, expense and time issues preclude such questioning for large-sample empirical ad studies. Hence, researchers typically rely on standardized measures of ethnic identity. Specifically, researchers assess ethnicity in two ways: self-designated ethnicity – identifying oneself as belonging to an ethnic group – and felt ethnicity – how strongly one identifies with an ethnic group. Self-designated ethnicity is measured with a single, closed-ended question that asks respondents to indicate if they are Hispanic, Black, White, Asian, or Other. Felt ethnicity or strength of ethnic identity often is measured with the multigroup ethnic identity measure (Phinney, 1992). This measure, which applies to diverse samples and has proven reliable in previous studies, consists of five statements about ethnic attachment, feelings about ethnic background, happiness with ethnicity, ethnic pride, and sense of ethnic belonging. Responses to each statement are captured by a seven-point Likert scale anchored by *strongly disagree* to *strongly agree*.

Review of Previous Empirical Studies on Ethnicity and Advertising

Studies that qualified for review examined advertising-related effects of ethnic identification. In most studies, ethnicity was manipulated as an independent variable. Typically, respondents first read a magazine with filler and test ads. The test ads were identical in all regards but model ethnicity. Then respondents indicated their attitudes and/or purchase intentions toward the target ads and brands (e.g., see ATTITUDES; ADVERTISING MESSAGE APPEALS; ADVERTISING EFFECTIVENESS; INTERNATIONAL ADVERTISING – IS THERE STILL A STANDARDIZATION VERSUS LOCAL ADAPTATION DEBATE?; INTERNATIONAL RETAILING). Many studies assessed ethnic identification with an ethnic-identity or perceived-similarity measure.

Studies were identified through (i) a keyword search of journal aggregator databases in busi-

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

ness, communications, education, and psychology (e.g., ABI/Inform, EBSCO, ERIC, PsycInfo), (ii) a search for conference proceedings in Papers First and marketing society web sites (e.g., AMA, SMA), (iii) an online search for articles using keywords from uncovered articles, and (iv) listserv requests to business and psychology researchers for copies of published or unpublished manuscripts. Excluded studies examined effects of ethnic identification on nonadvertising outcomes, such as political affiliation, psychological assessments, and educational outcomes. Because few studies examined brand-related outcomes, such as brand prestige, brand loyalty, and brand awareness, those constructs were ignored.

Ultimately, 25 empirical articles, all published since 1971, were identified for review. In terms of studies per decade, the 2000s (11) and 1990s (8) are most prolific. The most common publication outlets were *Journal of Advertising*, *Journal of Advertising Research*, *Journal of Marketing Research*, and *Psychology & Marketing*, with four articles each, respectively; and, the most prolific authors were Deshpandé (6 articles), Whittler (3 articles), and several researchers with 2 articles (e.g., Appiah, Brumbaugh). Table 1 summarizes the 25 articles.

Ethnicities Studied. The most common ethnic cues in test ads portray Whites (75% of studies), Blacks (57% of studies), Hispanics (21% of studies), and Asians (14% of studies). These four ethnic groups are the biggest spenders and represent the largest population subgroups in the United States (U.S. Census Bureau, 2001).

Theoretical Frameworks Applied. Nine theoretical frameworks have been used to explain the effects of ethnic identity in advertising. Each framework, which offers unique explanatory power, may be encapsulated as follows:

- *Accommodation theory*: proposes that people generally like other people who share similar traits, which suggests, for example, that Black viewers will respond favorably to ads that use Black actors.
- *Cultural script theory*: emphasizes the portrayal of cultural themes and values, distinct to an ethnic group, through social communication, which suggests that the use of Spanish or *Spanglish* verbiage in ads may resonant favorably with Hispanic viewers.
- *Distinctiveness theory*: posits that a person's distinctive or unique characteristics are more important to oneself than other local people's common traits. This theory suggests why Hispanics living in a Hispanic-minority-White-majority region are more likely to trust a Hispanic rather than White actor in an ad.
- *The elaboration likelihood model*: suggests that attitudinal responses stem from central route (i.e., high elaboration) or peripheral route (i.e., low elaboration) processes. As a result, Hispanic viewers may peripherally assess an ad for a Hispanic product when a Hispanic actor is used, and may centrally assess an ad for a Hispanic product when a White actor is used.
- *The heuristic–systematic persuasion model*: posits that message credence is evaluated either heuristically (i.e., casual evaluation) or systematically (i.e., scrutinized evaluation). For example, Asian viewers may systematically evaluate ads embedded with Asian cues because they trust an ethnically resonant source.
- *Identification theory*: suggests that people examine their similarity with environmental sources and then make similarity judgments, which may lead to Black viewers identifying more with ads that use Black actors rather than White actors.
- *In-group bias theory*: proposes that people favor in-group members, based on some characteristic, more than out-group members. As a result, Whites should evaluate White actors more favorably in ads with White and Asian actors.
- *Polarized appraisal theory*: suggests that in-group members will evaluate in-group stimuli less extremely than out-group stimuli. This theory implies that Hispanics will evaluate an ad with a Black spokesperson more thoroughly than an otherwise identical ad with a Hispanic spokesperson of comparable character.
- *Social identity theory*: posits that people's self-concept stems from their social and self-identities, which in turn contribute to self image and satisfaction. Hence, ads with

Table 1 Summary of ethnicity and advertising literature.

Article	*Research Question*	[a] *Overall Theory*	[b] *Ethnicity*	[c] *n*; [d] *Subjects*	[e] *Stimuli*	*Data Collected*	*Findings*
Muse (1971) (*JMR*)	How are ads with only Black models perceived by White audiences?	No framework used	Blacks, Whites	233; college students	Print ads for cigarettes, vodka, napkins, beer	Personal interviews; ads ranked on effectiveness and overall appeal	Ads with only Black models do not affect White consumers' perceptions of ad effectiveness or appeal
Bush, Gwinner, and Solomon (1974) (*JM*)	How do White consumers respond to Black models in in-store promotional materials?	No framework used	Blacks, Whites	13443; White adults from SW US	End-of-aisle, POP displays with either all Black, all White, or mix of Black and White models	Observers recorded shopping behavior in supermarket	Whites respond similarly to point-of-purchase (POP) displays with all Black models, all White models, or a mix of Black and White models
Deshpandé, Hoyer, and Donthu (1986) (*JCR*)	Do different groups of Hispanics differ in their ethnic identification?	No framework used	Hispanics	425; registered TX voters, 278 Whites, 147 Hispanics	None; mail questionnaire	Ethnicity; attitude toward institutions; A_{AD}; brand loyalty; purchase influences	Relative to weak Hispanic identifiers, strong Hispanic identifiers more likely (1) to use Spanish language media, (2) to have more positive attitudes toward advertising, and (3) to purchase products advertised to Hispanics
Pitts *et al.*, (1989) (*P&M*)	How do Blacks and Whites respond to TV ads with cultural/ethnic cues?	No framework used	Blacks	271; undergrads from W. and S. US; 82 Blacks, 189 Whites	Four 60-second TV ads targeted at Blacks	Personal values; ad rating; product usage rate	• Relative to Whites, Blacks respond more favorably to TV ads with Black actors

(*continued overleaf*)

Table 1 (*Continued*).

Article	*Research Question*	[a] *Overall Theory*	[b] *Ethnicity*	[c] *n*; [d] *Subjects*	[e] *Stimuli*	*Data Collected*	*Findings*
							• Compared to Blacks, Whites fail to notice some cultural values (e.g., belonging, self-fulfillment, accomplishment) depicted in commercials
Whittler (1989) (*P&M*)	Do race of ad actor and viewers' racial attitudes affect ad evaluations?	No framework used	Blacks, Whites	340; undergrads from US Midwest; 180 White, 160 Black	Full-color storyboard ads with a Black actor and a White actor	Message comprehension; racial attitudes	• Whites (Blacks) identify more with White (Black) actors than Black (White) actors • Blacks (Whites) are more likely to buy advertised brand and evaluate ads with a Black (White) actor more favorably • Whites do not react negatively to Black actors in ads
Webster (1992) (*JAR*)	Do Hispanic groups categorized by sub-cultural ethnic identification search differently for ad information?	No framework used	n/a	180; women from San Antonio, TX	None; administered questionnaire	Extent of information search in various media (e.g., ads, yellow pages)	• Hispanics more willing to search media with which they ethnically identify • Relative to non-Spanish-speaking Hispanics, Spanish-speaking Hispanics influenced less by brochure ads and magazine ads

Roslow and Nicholls (1996) (*JAR*)	Are Hispanics more persuaded by Spanish- or English-language TV ads?	No framework used	Hispanics	648; adult Hispanics from NYC, Miami, Houston, and LA	Spanish-language and English language TV ads	Ad persuasiveness	Hispanics more persuaded by Spanish-language TV ads embedded in Spanish-language TV programs than by analogous English-language TV ads embedded in English-language TV programs
Dimofte, Forehand, and Deshpandé (2004) (*JA*)	Can ad targeting incongruent with existing identity cues affect the salience of self-identification and responses of target consumers?	No framework used	Hispanics, Whites	82; employees of major California University	Two TV ads with a Hispanic actor; one ad with English voice-over, and one ad with Spanish voice-over and English subtitles	Ethnicity; ad schema congruity; A_{AD}; attitudes toward spokesperson; cognitive responses	• Unusual voice-over/sub-titling augments ethnic self-awareness and increases ad recall • Ad schema congruity moderates the effect of target market affiliation on A_{AD} and attitude toward spokesperson
Williams and Qualls (1989) (*P&M*)	Do middle class Blacks and Whites respond differently to ads featuring celebrity endorsers?	Cultural script	Blacks	160; adults from Denver CO, 80 Blacks, 80 Whites	TV ads featuring celebrity endorsers	Attitudes toward the spokesperson and product; A_{AD}	• Strong and weak Black identifiers respond similarly to ads with celebrity endorsers • Strong Black identifiers and Whites hold similar favorable attitudes toward ads with celebrity endorsers

(*continued overleaf*)

Table 1 (*Continued*).

Article	*Research Question*	[a]*Overall Theory*	[b]*Ethnicity*	[c]*n*; [d]*Subjects*	[e]*Stimuli*	*Data Collected*	*Findings*
Deshpandé and Stayman (1994) (*JMR*)	Do majority and minority group members respond differently to radio ads with White or Hispanic-named spokespersons?	Distinctiveness	Hispanics; Whites	205; adults from San Antonio and Austin, TX	Radio-script ads with either a White or Hispanic-named spokesperson	Ethnic salience; spokesperson trustworthiness; brand attitudes	Relative to majority group members, minority group members more likely (1) to deem their ethnicity important, and (2) to trust spokespersons of similar ethnicity, which induces more positive A_B
Grier and Brumbaugh (1999) (*JA*)	What meanings do targeted and nontargeted groups derive from ads?	Distinctiveness	Blacks, Whites, gay/lesbian	62; MBAs; 35 White, 20 Black, 7 unknown; 40 straight, 22 other	Ad brochure that mimics magazine ads	Thoughts and feelings about the ad	Relative to non-targeted groups, targeted groups (1) view ads more positively, and (2) better understand cultural cues in ads
Aaker, Brumbaugh, and Grier (2000) (*JCP*)	What effect do ads targeted at one audience have on non-targeted audiences?	Distinctiveness	Blacks, Whites, gay/lesbian	63; MBAs from US Midwest; 123; students, 39 Black, 84 White	Pamphlet of color print ads targeting Blacks, Whites, or gay/lesbians	A_{AD}; felt targetedness	• Relative to targeted group, nontargeted groups' view ads less positively • When targeting cues absent, viewer distinctiveness alone unrelated to A_{AD} • Felt similarity and felt targetedness mediate A_{AD} for intended targets

Forehand and Deshpandé (2001) (*JMR*)	Does ethnic self-awareness affect consumers' responses to targeted ads?	Distinctiveness	Asians, Whites	109 and 175; Asian and White undergrads from W. US	Series of TV and print ads	Attitude toward spokesperson; A_{AD}; strength of ethnic identity;	When ethnic cues are analogous to one's self-concept, they increase the rate at which people mention their ethnicity in self-descriptions, leading to more favorable responses to a same-ethnicity spokesperson and ads targeting their ethnicity
Grier and Deshpandé (2001) (*JMR*)	Does social and numeric status of consumers affect ad effectiveness?	Distinctiveness	Blacks, Whites	176; S. African women	One print ad with either a Black or White spokesperson	Racial/ethnic salience and brand attitude	• Using social dimensions to target ads may be effective even when targeted group is a numeric majority • Greater (lesser) ethnic salience due to minority (majority) status leads to more (less) positive A_B
Forehand, Deshpandé, and Reed (2002) (*Journal of Applied Psychology*)	How do ethnic primes and social distinctiveness affect identity salience and responses to targeted ads?	Distinctiveness	Whites, Asians	284 and 384; Asian and White undergrads from W. US	Experiment 1: Four TV ads, 1 for Nokia and 3 filler ads	Spokesperson liking; cognitive responses; A_{AD}	Across both experiments, Asians (Whites) responded more favorably (negatively) to an Asian spokesperson and Asian-related ads when they were both primed and socially distinctive (i.e., numeric minority group in a social environment)

(*continued overleaf*)

Table 1 (*Continued*).

Article	*Research Question*	[a]*Overall Theory*	[b]*Ethnicity*	[c]*n*; [d]*Subjects*	[e]*Stimuli*	*Data Collected*	*Findings*
					Experiment 2: 2 print ads for Northwest Airlines		
Lee, Fernandez, and Martin (2002) (*International Journal of Advertising*)	How do ethnic minority models in ads affect the evaluations of ethnic minority and ethnic majority consumers?	Distinctiveness	Asians, Whites	178; college students	Full-color print ads for watches and facial tissue with either Asian or White models	Self-referencing; ethnic identity; cognitive responses; A_{AD}; A_B; A_M	Consumers exposed to ads consistent with their ethnicity, spontaneously self-reference the ad, which leads to (a) more positive attitudes toward the ad and the model depicted in the ad, (b) stronger intentions to buy the advertised brand, and (c) more favorable impressions of the advertised brand
Martin, Lee, and Yang (2004) (*JA*)	Does consumer self-referencing mediate the effect of ethnicity on ad/brand-related attitudes and intentions to buy the advertised brand?	Distinctiveness	Asians, Whites	122 college students; 66 Whites, 56 Asians	Tea and shampoo print ads with White and Asian female models	Self-referencing; A_{AD}; A_B; A_M	• Self-referencing mediates ethnicity effects on (a) ad-related attitudes, and (b) PI_B • For Asians, self-referencing relates positively to A_{AD} and PI_B • For Asians and Whites, Asian models in ads for atypical products boosted (a) self-referencing, (b) ad-related attitudes, and (c) PI_B

							• Relative to Whites, Asians exhibit greater self-referencing of Asian models in ads
Appiah (2001b) (*JAR*)	Does strength of ethnic identity affect Black and White adolescents' responses to print ads with models of different races?	Distinctiveness/ Identification	Blacks, Whites	173; high school students from CA, 81 Blacks, 92 Whites	Three full-color print ads, with race of characters and cultural cues digitally manipulated	Perceived similarity with actors; belief ad intended for subject; identification with actors; A_{AD}	• Relative to weak Black identifiers, strong Black identifiers see themselves as more similar to and identity more strongly with Black characters in ads • Whites' identification with characters or A_{AD} is unrelated to strength of ethnic identity
Appiah (2001a) (*Howard Journal of Communications*)	How do adolescents of different races respond to ads with Black or White actors?	Distinctiveness/ Identification	Blacks, Whites	349; high school students from CA, 81 Blacks, 84 Asian, 92 Hispanics, 92 Whites	Four full-color print ads, with Black or White characters, for consumer non-durables	Perceived similarity and identification with characters; belief ad intended for subject; ad rating	• Blacks more influenced by race of ad actor than ethnic cues in ad • Identification with an ad by Hispanics, Blacks, and Asians influenced by ad actor's Race • Blacks, Whites, Hispanics, and Asians rate Black-actor ads more favorably than White-actor ads

(*continued overleaf*)

Table 1 (*Continued*).

Article	*Research Question*	[a]*Overall Theory*	[b]*Ethnicity*	[c]*n*; [d]*Subjects*	[e]*Stimuli*	*Data Collected*	*Findings*
Koslow, Shamdasani, and Touchstone (1994) (*JCR*)	How do Hispanics respond to ethnic language in ads?	Accommodation	Hispanics, Whites	367; Spanish speakers	Four print ads: 1 English only, 1 Spanish only, and 2 mixed English and Spanish	A_{AD}; advertiser's sensitivity to Hispanic culture; ethnic identity	Spanish-language ads enhance Hispanics' beliefs about advertiser's sensitivity to Hispanic culture, which boosts Hispanics' A_{AD}
Green (1999) (*JA*)	Are ethnic identity, media placement, and race of ad actor(s) related to Blacks' A_{AD} and PI_B?	Accommodation/in-Group bias	Blacks, Whites	313; Black American females	Full-color print ads, with either White or Black models, for a fictitious brand	Ethnic identity; $A_{AD;}$ PI_B	Relative to weak ethnic identifiers, strong ethnic identifiers (1) evaluate ads that feature Blacks more positively, and (2) have stronger purchase intentions for brands with ads that feature Blacks
Qualls and Moore (1990) (*P&M*)	Does race affect Blacks' and Whites' ad evaluations?	In-group bias/ polarized appraisal	Blacks, Whites	211; beer drinkers from Midwest town in US; 103 Blacks, 107 Whites	Experimental TV ad for a new beer	A_{AD}; A_B; A_M	• Relative to polarized appraisal theory, in-group bias theory better explains effect of race on people's ad evaluations • Product evaluations by Blacks (Whites) improve (worsen) for ads with Black actors and worsen (improve) for ads with White actors

Sierra, Hyman, and Torres (2009) (*Journal of Current Issues and Research in Advertising*)	What effect does ethnic identification with print ads, created by the model's apparent ethnicity, have on A_{AD}, A_B, A_M, and PI_B?	Social Identity	Whites, Blacks, Hispanics	207; undergrads from SW town in US; 93 Whites, 62 Blacks, 52 Hispanics	Full-color print ads, with White, Black, or Hispanic male models, for a fictitious athletic shoe brand	Ethnic identification with ad; A_{AD}; A_B; PI_B	Ethnically identifying with print ad, based on model's apparent ethnicity, results in positive direct and indirect (through A_{AD} and A_B) effect on PI_B
Whittler and DiMeo (1991) (*JAR*)	What effect does viewer's prejudice have on simple decision rules when examining ad effectiveness?	Heuristic-Systematic Persuasion Model	Blacks, Whites	160; paid volunteers	Full-color story-board ads for a fur coat and laundry detergent	Brand awareness and PI_B; ad recall and impression; likability of, and perceived similarity with actor; racial attitudes; message claim validity	• Low-prejudice Whites (a) believe equally similar to White and Black actors, and (b) identify equally with White and Black actors • High-prejudice Whites (a) believe less similar to Black than White actors, and (b) identify more strongly with White than Black actors • Regardless of their attitudes toward Blacks, Whites had lower PI_B and less favorable A_{AD} and A_B for ads featuring a Black actor

(*continued overleaf*)

Table 1 (*Continued*).

Article	*Research Question*	[a] *Overall Theory*	[b] *Ethnicity*	[c] *n*; [d] *Subjects*	[e] *Stimuli*	*Data Collected*	*Findings*
Whittler and Spira (2002) (*JCP*)	What effect does race in advertising have on viewer message processing and does it function as a peripheral cue?	Elaboration Likelihood Model	Blacks, Whites	160 Black adults from the SE US	Full-color storyboard ads for cordless phone, garment bag, and laundry detergent	Same as Whittler and DiMeo (1991) plus message comprehension and identification with Black culture	• Identification with Black culture moderates Blacks' responses to ads • High-identification Blacks respond more favorably to ads, products, and models when the ad model is Black • Ad model's race influences Blacks' peripheral message processing • Blacks think more about products in ads with White rather than Black models • Blacks better recall ads with Black models

[a] general theoretical framework that grounded study
[b] ethnicity depicted in the test ads
[c] sample size for each study; note that some articles described more than one study
[d] study participants
[e] ads or commercials that participants read or viewed
Note: *JA = Journal of Advertising; JAR = Journal of Advertising Research; JCP = Journal of Consumer Psychology; JCR = Journal of Consumer Research; JM = Journal of Marketing; JMR = Journal of Marketing Research; P&M = Psychology & Marketing;* A_{AD} = attitude toward the ad; A_B = attitude toward the brand; A_M = attitude toward the model; PI_B = intentions to buy the advertised brand

embedded ethnic cues allow viewers in the targeted ethnic group to differentiate themselves from others, which reinforces their self-identity and uniqueness.

Collectively, these theoretical frameworks suggest that favorable advertising effects may ensue from ethnic identification with ads. They offer insights about greater responsiveness to ads that reflect resonant cultural values, increased trustworthiness of actors of similar ethnicity, and enhanced attitudes toward ads that imply cultural sensitivity to ethnic minorities; yet, researchers' reliance on these various frameworks does not explain the inconsistent findings on ethnic-identity effects in advertising. The frameworks most often applied to ethnic identity studies are distinctiveness theory (10 studies) and two studies each relied on identification theory, in-group bias theory, accommodation theory, and the heuristic-systematic persuasion model.

Standard Study Procedures. The most recurrent research question was response differences among ethnic groups – Blacks, Whites, Hispanics, and Asians – to ads with various ethnic cues. At least one ethnic cue appears in the test ads for all but two studies. Test ads depicted only two ethnic groups in 80% of studies, which may inhibit the generalizability of reported findings (Brumbaugh and Grier, 2006). Print or television ads were assessed in 92% of studies. The most common ethnic cues in test ads are targeted at Whites (77% of studies), Blacks (54% of studies), and Hispanics (22% of studies).

The mean sample size, which ranges from 62 to 648 (excluding an ethnographic study with 13 443 observations), is 239 people. The samples represent much of the United States where eight are Western, five are Southwestern, and three are Midwestern and Southern. Eleven studies rely on student samples – seven of undergraduate students, two of MBA students, and two of high school students – and eleven studies rely on adult samples. Thus, reported findings reflect responses of a roughly even mix of students and adults located throughout the United States.

In terms of the methods employed, typically, respondents were first exposed to advertising stimuli, including filler and test ads. For the most part, test ads were identical in all regards but model ethnicity. Subsequently, responses were gathered about the ad, advertised brand, and/or model. Data pertaining to ethnic identity or perceived similarity with the model were often collected. Regarding the theoretical constructs examined, all studies assess some consumer attitudes (e.g., attitude toward the ad, attitude toward the brand), but only 10 studies measure intention to buy the advertised brand and only 3 studies measure ethnic identification with the ad. Only 1 study examines ethnic identification with the ad and its effect on purchase intentions of the advertised brand.

The most frequently used statistical method is ANOVA, which appears 13 times; the other methods used are regression analysis (7 times), MANOVA (4 times), discriminant analysis (2 times), and path analysis (2 times). Although 15 studies were published recently, 1994 or later, structural equation modeling is used only once.

Findings. The findings generally suggest that non-White consumers, relative to White consumers, are aware of and respond more favorably to ads with ethnic cues. In five studies that contrast weak versus strong ethnic identifiers, same-ethnicity actors in ads are more favored by the latter group. The four studies on the effect of ads with ethnic language (e.g., ads with Spanish words embedded) on same-ethnicity consumers suggest that Hispanics are more likely (i) to use Spanish-language media, (ii) to have more positive attitudes toward ads with Spanish verbiage, and (iii) to buy a product advertised with Spanish verbiage.

In-group members of an ethnic group evaluate ad stimuli more favorably when the actors/models are from the in-group rather than the out-group (Qualls and Moore, 1990). For example, relative to Whites, Blacks develop more positive affect toward ads with Black actors (Pitts *et al.*, 1989). Whites (Blacks) identify more with White (Black) actors than Black (White) actors, respond more favorably to ads with White (Black) actors, and are more likely to buy the advertised brand when White (Black) actors are featured (Whittler, 1989). Hence, ethnic consumers should identify with and respond more favorably to ads that target their

ethnicity. However, such assumptions are not always supported. Research shows that Whites respond similarly to point-of-purchase displays with all Black models, all White models, or an amalgamation of White and Black models (Bush, Gwinner, and Solomon, 1974). Using Asian models in ads favorably affects attitudes and purchase intentions toward the featured brand for both Asian and White consumers (Martin, Lee, and Yang, 2004). Asians, Blacks, and Hispanics identify most with ads that use Black actors, and Asians, Blacks, Hispanics, and Whites evaluate ads that use Black actors more positively than ads featuring White actors (Appiah, 2001b). These inconsistent findings cloud the role of ethnic identification in advertising and complicate decisions about ad design.

Are Study Results Artifactual? Although most ethnic identification studies relied on test ads evaluated by subjects in experimental or quasi-experimental settings, differences in method and research context may affect variance homogeneity across studies. Four possible sources of artifactual results are measurement type (i.e., single-item measure vs multi-item scale), study type (i.e., survey-based vs experiment-based), sample type (e.g., student or random adult), and media type (i.e., print vs TV).

In many cases, single-item measures can provide good assessments of concrete and well-accepted constructs, such as immediate purchase intention. However, single-item measures tend to reflect attitudinal and complex constructs less reliably, and lower reliability may cause less stable correlations. Conversely, typically more reliable and valid multi-item scales should yield more stable correlations.

Relative to surveys, experiments – by permitted control of extraneous factors that otherwise would interact with studied constructs – can generate larger effect sizes and imply inter-construct relationships that ultimately prove artifactual. Although the potentially interactive extraneous factors are infinite, many of them become known as a research stream matures.

Heterogeneous samples tend to attenuate effect size, especially in experimental studies (Fern and Monroe, 1996). Conversely, homogeneous samples may produce restricted responses due to similar respondent backgrounds, incomplete self-identity, strong need for approval, and unstable group relationships. Although some marketing scholars have found that students respond more homogeneously, and other social scientists have found limited evidence that student samples yield larger effect sizes, a recent comprehensive meta-analysis suggests otherwise (Peterson, 2001).

Although previous research showed that television ads are more effective than print ads, recent studies suggest that television ads may only dominate print ads on ad awareness (Hansen, Olsen, and Lundsteen, 2006). Perhaps ethnic-identity effects are larger for television ads than for print ads because television's multichannel visual and auditory cues induce higher-quality thoughts and feelings for ethnically resonant ads.

The large number of theoretical frameworks and small number of published studies limited framework comparisons to no-stated framework, distinctiveness theory, and in-group bias theory. Studies with no-stated framework reported significantly higher effect sizes than in-group bias studies, which in turn reported higher effect sizes than distinctiveness theory studies. Type of measurement (single vs multi-item outcome scale) and study design (experiment vs survey) help to explain some heterogeneity in attitude toward the ad, model, brand, and purchase intentions of the advertised brand. As expected, ethnic identity effect sizes were larger in experiment-based than survey-based studies. Studies with respondents from only two ethnic groups (e.g., Hispanic and Caucasian or Black and Caucasian) produced lower effect sizes than studies with more diverse samples of three or more ethnic groups. Surprisingly, single-item measures yielded larger correlations than multi-item scales, perhaps due to the larger variety of multi-item scales used in ethnic identity studies.

The sample-weighted estimates of ethnic identity's influence on all advertising outcomes were medium-sized effects with composite r ranging from 0.18 to 0.30. Five studies reported small effect sizes ($r < 0.10$) and none reported large effect sizes ($r > 0.50$). All correlations between ethnic identity and attitude toward the ad, attitude toward the model, attitude toward

the brand, and brand purchase intention were significant ($p < 0.05$).

Although methodological heterogeneity accounts for some unexplained variance in ethnic identity research, the relatively few studies with significant findings makes this result tenuous. Although researchers continue to debate the merits of student samples, we found no differences in studies with student versus nonstudent respondents. In the many experiment-based studies, researchers were able to maintain high internal validity and boost statistical power. As expected, questionnaire-based studies produced lower ethnic identity effect sizes. Although it does not account for response differences among ethnic groups, our analysis confirms that ethnic identity effects are activated in ethnic groups even if they are not portrayed in test ads. Thus, generalizability concerns about ethnic identity studies in advertising may include both the limited range of test stimuli (Brumbaugh and Grier, 2006) and respondent samples.

Discussion

Consumers relate to ethnically resonant ads by developing, maintaining, and enhancing their social identification and attachments to the featured brands. The results from our review of the literature indicate that ethnic identity influences all elements of the advertising hierarchy-of-effects, producing the largest effect on attitude toward the brand. The regression analysis of unique variances shows a direct linkage between attitude toward the model and attitude toward the brand. The moderation analysis suggests methodological artifacts; for example, ethnic identity effects are larger in experiment-based than survey-based studies. In addition, three studies with single-item measures produced higher correlations than studies with multi-item scales, especially with attitude toward the ad. Two of these studies examined attitude only toward the ad, and thus may have captured overall attitudes within a single-item attitude toward the ad measure.

Managerial Implications. Research shows that consumers are more likely to notice and respond favorably to an ad if it contains ethnically resonant cues (e.g., Qualls and Moore, 1990; Whittler, 1989); thus, advertisers should benefit from developing ads with such cues. Advertisers could accentuate same-ethnicity models' skin color, facial features, demeanor, verbal expression, and apparel styles, like McDonald's has done with its successful Hip Hop *i'm lovin' it* ad campaign targeted at Hispanic and Black customers. With their growing popularity and efficacy, the physical traits of animated spokescharacters (e.g., skin color, hairstyles, facial features) could be adapted to targeted ethnic viewers. Ethnic identity expands beyond actors; hence, durable goods producers can foster ethnic identification via ads that focus on country of origin. For example, ads targeting Asians can stress that some product components were produced and/or the product was assembled in China. Moreover, firms can appeal to Hispanics with ads that combine Spanish and English (i.e., Spanglish) without fear of backlash from Whites (Koslow, Shamdasani, and Touchstone, 1994).

Service providers can foster ethnic identification within targeted groups via ads that celebrate annual ethnic-related occasions, such as Saint Patrick's Day (e.g., Bennigan's Grill & Tavern offering green lagers), Cinco de Mayo (e.g., Chili's Bar & Grill promoting Mexican beer), and Black History Month (e.g., Barnes & Noble bundling Black literature). To appeal to Hispanics, firms could place ads with Hispanic actors in magazines targeting Whites. Although firms may believe that Hispanic actors in ads can induce negative attitude toward the ad and brand in White readers (Qualls and Moore, 1990), previous research shows that this belief is false (Whittler, 1989). Hence, ads with Hispanic actors will gain the attention of Hispanic readers without alienating White readers.

Future Research. Because ethnic identity significantly influences all stages of the hierarchy-of-effects model, future studies should explore the indirect and mediation pathways between advertising and ethnic identity, particularly within different advertising contexts with various products. Also, the relative efficacy of different types of ad cues (e.g., slogans, argot) to induce ethnic identification is unknown. Although our meta-analysis uncovered moderating factors that explain some variance in ethnic identity effects, further research is needed to answer questions like "Does the tagline or

actor's appearance in an ad create stronger ethnic identity effects?"

Bibliography

Aaker, J.L., Brumbaugh, A.M., and Grier, S.A. (2000) Nontarget markets and viewer distinctiveness: the impact of target marketing on advertising attitudes. *Journal of Consumer Psychology*, **9** (3), 127–140.

Appiah, O. (2001a) Black, White, Hispanic, and Asian American adolescents' responses to culturally embedded ads. *Howard Journal of Communications*, **12** (1), 29–48.

Appiah, O. (2001b) Ethnic identification on adolescents' evaluations of advertisements. *Journal of Advertising Research*, **41** (5), 7–22.

Brumbaugh, A.M. and Grier, S.A. (2006) Insights from a 'failed' experiment: directions for pluralistic, multi-ethnic advertising research. *Journal of Advertising*, **35** (3), 35–46.

Bush, R.R., Gwinner, R.F., and Solomon, P.J. (1974) White consumer sales response to Black models. *Journal of Marketing*, **38** (2), 25–29.

Deshpandé, R., Hoyer, W.D., and Donthu, N. (1986) The intensity of ethnic affiliation: a study of sociology of ethnic consumption. *Journal of Consumer Research*, **13** (2), 214–220.

Deshpandé, R. and Stayman, D.M. (1994) A tale of two cities: distinctiveness theory and advertising effectiveness. *Journal of Marketing Research*, **31** (1), 57–64.

Dimofte, C.V., Forehand, M.R., and Deshpandé, R. (2004) Ad schema incongruity as elicitor of ethnic self-awareness and differential advertising response. *Journal of Advertising*, **32** (4), 7–17.

Fern, E.F. and Monroe, K.B. (1996) Effect-size estimates: issues and problems in interpretation. *Journal of Consumer Research*, **23** (2), 89–105.

Forehand, M.R. and Deshpandé, R. (2001) What we see makes us who we are: priming ethnic self-awareness and advertising response. *Journal of Marketing Research*, **38** (3), 336–348.

Forehand, M.R., Deshpandé, R., and Reed, A. (2002) *II* Identity salience and the influence of differential activation of the social self-schema on advertising response. *Journal of Applied Psychology*, **87** (6), 1086–1099.

Green, C.L. (1999) Ethnic evaluations of advertising: interaction effects of strength of ethnic identification, media placement, and degree of racial composition. *Journal of Advertising*, **28** (1), 49–64.

Grier, S.A. and Brumbaugh, A.M. (1999) Noticing cultural differences: ad meanings created by target and non-target markets. *Journal of Advertising*, **28** (1), 79–93.

Grier, S.A. and Deshpandé, R. (2001) Social dimensions of consumer distinctiveness: the influence of social status of group identity and advertising persuasion. *Journal of Marketing Research*, **38** (2), 216–224.

Hansen, F., Olsen, J.K., and Lundsteen, S. (2006) The effects of TV vs. print advertising, documented using STAS measures. *International Journal of Advertising*, **25** (4), 431–446.

Koslow, S., Shamdasani, P.N., and Touchstone, E.E. (1994) Exploring language effects in ethnic advertising: a sociolinguistic perspective. *Journal of Consumer Research*, **20** (4), 575–585.

Lee, C.K.-C., Fernandez, N., and Martin, B.A.S. (2002) Using self-referencing to explain the effectiveness of ethnic minority models in advertising. *International Journal of Advertising*, **21** (3), 367–379.

Martin, B.A.S., Lee, C.K., and Yang, F. (2004) The influence of ad model ethnicity and self-referencing on attitudes. *Journal of Advertising*, **33** (4), 27–37.

Muse, W.V. (1971) Product-related response to use of Black models in advertising. *Journal of Marketing Research*, **8** (1), 107–109.

Peterson, R.A. (2001) On the use of college students in social science research: insights from a second order meta-analysis. *Journal of Consumer Research*, **28** (3), 450–461.

Phinney, J.S. (1992) The multigroup ethnic identity measure – a new scale for use with diverse groups. *Journal of Adolescent Research*, **7** (2), 156–176.

Pitts, R.E., Whalen, J.D., O'Keefe, R., and Murray, V. (1989) Black and White response to culturally targeted television commercials: a value-based approach. *Psychology and Marketing*, **6** (4), 311–328.

Qualls, W.J. and Moore, D.J. (1990) Stereotyping effects on consumers' evaluation of advertising: impact of racial differences between actors and viewers. *Psychology and Marketing*, **7** (2), 135–151.

Roslow, P. and Nicholls, J.A.F. (1996) Targeting the Hispanic market: comparative persuasion of TV commercials in Spanish and English. *Journal of Advertising Research*, **36** (3), 67–77.

Sierra, J.J., Hyman, M.R., and Torres, I.M. (2009) Using a model's apparent ethnicity to influence viewer responses to print ads: a social identity theory perspective. *Journal of Current Issues and Research in Advertising*, **31** (2), 41–66.

Tajfel, H. (1978) *The Social Psychology of Minorities*, Minority Rights Group, New York.

U.S. Census Bureau (2001) National population estimates. June, 18.

Webster, C. (1992) The effects of Hispanic subcultural identification on information search behavior. *Journal of Advertising Research*, **32** (5), 54–62.

Whittler, T.E. (1989) Viewers' processing of actor's race and message claims in advertising stimuli. *Psychology and Marketing*, **6** (4), 287–309.

Whittler, T.E. and DiMeo, J. (1991) Viewers' reactions to racial cues in advertising stimuli. *Journal of Advertising Research*, **31** (6), 37–46.

Whittler, T.E. and Spira, J.S. (2002) Model's race: a peripheral cue in advertising messages?. *Journal of Consumer Psychology*, **12** (4), 291–301.

Williams, J.D. and Qualls, W.J. (1989) Middle-class Black consumers and intensity of ethnic identification. *Psychology and Marketing*, **6** (4), 263–286.

globally integrated marketing communications

Shintaro Okazaki

Definition and Conceptualization

The concept of globally integrated marketing communications (GIMC) was first introduced by Okazaki, Taylor, and Zou (2006) in their work on advertising standardization strategy by American and Japanese multinationalcorporations (MNCs) in the European Union. They refer to GIMC in the traditional sense of communicating a single message through a multiple marketing communication mix in various media across countries, and conceptualized it as a key element of global strategic orientation.

However, the ability to engage in GIMC depends not only on the firms' strategic orientation, but also on the availability of agencies that operate on a global basis. In this regard, the literature suggests that the growth of "mega-agencies" has changed the structure and operational practices of IMC. Most MNCs now seek simple but self-consistent solutions for global planning and implementation facilities in multiple markets. Agencies also believe that the full-service advertising agency will remain the principal source of creative and media services (Tharp and Jeong, 2001). The future use of standardization is more likely to be increased by these internal and external pressures than by economic benefits or the emergence of global consumers (Duncan and Ramaprasad, 1995). However, Gould, Lerman, and Grein (1999) point out that large, US-based advertising agencies are likely to assume that "the main issue of global communications is not thought to be standardization versus adaptation, per se, but rather organizational coordination which recognizes and encourages global strategies while working with local managers and markets" (see page 13). Such integration seems possible only for those "mega-agencies" that hold a wide variety of international advertising accounts in diverse regions (Moriarty and Duncan, 1990). In addition, the proliferation of the Inter/intranet, the digitalization of traditional media, and mergers and acquisitions among media companies have enabled agencies to practice media planning on a global basis (Okazaki, 2005).

In this vein, Tharp and Jeong (2001) provide an astute observation on global network communications agencies. They argue that the advertising agency's perspective is very relevant to the global integration of marketing communications, because it is unlikely that MNCs make the very technical decisions such as the adoption and planning of, for example, global websites. Tharp and Jeong argue that "Global clients want 'one-stop access' and global planning and implementation capabilities from their agency partners", and these agencies are particularly helpful in predicting consumer characteristics in new markets and new media. In addition, larger agencies are expanding their service portfolios horizontally into areas requiring more specialized expertise, by merging interactive advertising, strategic brand planning, and direct marketing (Tharp and Jeong, 2001). This implies that such large agency networks may have consolidated websites in different parts of the world. Therefore, the presence in Europe of US-based multinational agencies, such as Omnicom Group (which includes *BBDO*, *DDB Needham*, *TBWA*, etc.) and Interpublic Group (which includes *Ammirati Puris Lintas*, *The Lowe Group*, *McCann-Erickson*, etc.) may have substantially influenced the GIMC of American MNCs.

More evidence of the emergent global mega-agency can be found in the literature on the standardization of corporate visual identity from a UK perspective. Melewar and Saunders (1999) examined the factors influencing the international standardization of corporate visual identity among 40 British MNCs that operate in Malaysia. They found that local agencies play a minimal role in the development of corporate visual identity for the firms. Instead, British MNCs tend to use UK-based agencies or the in-house designers at their corporate head office. This appears to support our thesis in that the larger firms that operate in international markets centralize their marketing communication strategy, which is consistent with our conceptualization of GIMC.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

The Primary Benefits of GIMC

The primary benefits of GIMC can be summarized as follows:

1. The creation of a uniform brand image across markets.
2. Lower costs in developing and maintaining marketing strategy and programs.
3. The execution of common campaigns to cross-border segments that share similar preferences.

Some may argue that the adoption of GIMC is still limited in the sense that the integration of the mix of various marketing communication strategies and programs is a rather complicated task (Kitchen and Schultz, 1999). For example, the most visible element of marketing communication, advertising, is relatively easier to coordinate across borders, compared with other elements, such as sales promotion or direct marketing. Nevertheless, GIMC is an increasingly important issue, as it is directly related to the creation of a strong and uniform brand image across markets. By establishing consistent logos and symbols, sponsorship, events and experiences, and Internet website, the brand image – especially in terms of awareness, uniqueness, and association – can be similarly perceived in multiple markets. Undoubtedly, this leads to the strengthening of the brand equity.

In addition, the worldwide coordination of marketing communication would decrease the production costs, because relevant information about implementation and expertise can be shared among foreign subsidiaries and headquarter(s) via the networks of global mega-agencies (Tharp and Jeong, 2001). In the same token, the costs would be lower when they share advertising campaign or commerce globally (Zou and Cavusgil, 1996). Currently, the prolonged worldwide economic recession has made it imperative that firms seek a more cost-effective communication strategy as many companies are cutting their marketing communications budget (Goldsmith, 2009). The digitalization of media planning and execution is making this coordination task more realistic and feasible than it was in the past (Okazaki, 2005).

Finally, the availability of integrated media enables firms to seek commonly attractive and persuasive campaigns to cross-border segments. The literature suggests that there is a global youth segment that shares similar lifestyles, habits, and preferences in many industrialized countries. For example, the entertainment industry – especially gaming, is an example of a cross-border segment that can employ similar promotional campaigns with almost identical stimuli and messages.

Theoretical Perspectives

Standardization of global marketing. One of the reasons why firms should integrate their marketing communication mix across borders is related to the global marketing strategy (GMS) of international firms. In this regard, Zou and Cavusgil (2002) presented a broad conceptualization of a theory of GMS that incorporates eight dimensions, including product standardization, promotion standardization, distribution standardization, pricing standardization, and other dimensions related to the global coordination of value-adding activities. Building on industrial organization theory and a resource-based view of the firm, the GMS theory contends that the fit between (i) a company's marketing strategy and (ii) its external environment and its internal organizational resources determines the company's performance in the global market. When external market factors and internal organizational characteristics are conducive to global marketing, a more globalized marketing strategy, such as employing a higher degree of standardization and integration, will positively impact a company's strategic and financial performance globally. Using survey data, Zou and Cavusgil (2002) offered empirical support for the GMS model.

Among the various strategies of GIMC, probably one of the most visible examples is the standardized advertising of global brands. For over 50 years, the topic of standardization of advertising has received considerable attention in academic literature, and it has developed into one of the most researched topics in international advertising (Agarwal, 1995). The issue of standardization arises from the concept of global marketing which notes the desirability and feasibility of the use of a uniform marketing

mix (4Ps – product, price, place, and promotion) across international markets (Szymanski, Bharadwaj, and Varadarajan, 1993). Among them, advertising has been examined more often than the other elements. A "standardized" approach is the use of uniform messages with no modification of headings, graphics, or copy, except for translation, in international markets (Onkvisit and Shaw, 1987). This school of thought argues that consumers anywhere in the world are likely to share the same wants and needs (Elinder, 1961; Levitt, 1983). On the other hand, a "localized" approach asserts that consumers differ from country to country, and advertising should thus be tailored according to the culture, media availability, product life cycle stages, and industry structure (Synodinos, Keown, and Jacobs, 1989). Between these two extremes, a third school of thought states "think global/act local", as the appropriateness of standardization depends on economic similarity, market position, nature of the product, environment, and organizational factors (Jain, 1989).

In the 1970s, empirical evidence indicated a high degree of localization owing to increasing nationalistic forces as well as well-publicized advertising blunders during the 1960s (Agrawal, 1995). This trend reversed toward standardization in the 1980s along with the rapid rise of multinational advertising agencies. During this period, a series of content analysis studies attempted to identify cross-cultural differences between Japanese and American advertising (Belk, Bryce, and Pollay, 1985; Mueller, 1987).

In the 1990s, localization appears to have remained popular among MNCs operating in the various regions of world markets. Harris (1994) found that 69% of 38 MNCs (19 Americans and 19 Europeans) standardized their advertising campaigns to some extent in European markets, while the rest of the sample localized. However, only 8% of the sample used totally standardized advertising, providing "little evidence of any widespread practice of standardized pan-European advertising campaigns" (Harris, 1994). Kanso and Nelson (2002) found that 62% of 193 firms studied (both American and non-American subsidiaries) in Finland and Sweden used localization, placing strong emphasis on local cultures. Similarly, Samiee *et al.* (2003) found that MNCs operating in Southeast Asia tend to localize advertising. They examined 113 firms in Hong Kong, the Peoples' Republic of China, Taiwan, and Singapore, and found that environmental and economic factors are the primary drivers of this tendency.

In a recent study, Okazaki, Taylor, and Zou (2006) proposed a model of advertising standardization consisting of organizational, environmental, and strategic factors as antecedents, and advertising effectiveness, financial performance, and strategic performance as consequences. Given that advertising standardization is a component of the GMS, they applied Zou and Cavusgil's (2002) model, which provides a solid theoretical foundation for investigating the drivers and consequences of advertising standardization. On the basis of the data collected in several European Union member states (United Kingdom, France, Spain, Germany, and the Netherlands), their findings suggest that the recent economic unification in Europe has contributed to the acceleration of cross-market homogeneity in consumer characteristics, market structure, media availability, and competitive intensity, which, in turn, closely relates to the degree of standardization of advertising and its effectiveness.

Global brand positioning theory. Another relevant theoretical framework for GIMC is global brand positioning theory. Drawing on global consumer culture theory, Alden, Steenkamp, and Batra (1999) have suggested that "global consumer culture positioning" (GCCP) will resonate with increasingly global segments of consumers. The basic idea is that a gradual trend toward cultural convergence in some (but not all) respects makes standardization more feasible than in the past, but that some types of marketing approaches can be standardized more easily than others. In GCCP, the brand is defined as *a symbol of a given global culture, which consumers may purchase to reinforce their membership in that segment*. Obviously, from a GIMC perspective, the existence and prevalence of a global consumer culture is a good reason for firms to integrate, concentrate, and coordinate their marketing communication mix across countries.

Specifically, Alden, Steenkamp, and Batra (1999) argue that emotional and image-based

advertising messages or a "soft-sell approach" is more suitable than hard-hitting and fact-based advertising appeals or "hard-sell approach" for GCCP, because "global consumer culture is an emerging and rapidly changing phenomenon, with differing sets of signs in differing global segments. Advertising that uses this "soft-sell" approach will be more effective because it communicates in a subtle, indirect, and abstract fashion" (Alden, Steenkamp, and Batra, 1999, p. 79). The authors classified the overall sales appeal of an advertisement by labeling two contrasting approaches: the "soft-sell" approach (image-oriented content that does not emphasize reasons to buy, but rather conveys general associations with the brand) versus the "hard-sell" or direct approach. It is of interest to note that their research suggested that the features of the "soft-sell" approach (subtlety, implicitness, and abstractness) are more suitable than the "hard-sell" approach for a global consumer culture positioning strategy.

In the study, the authors analyzed a total of 1267 advertisements from seven countries. In total, 284, or 22.4%, of the advertisements were found to employ a GCCP. More than half (56.4%) of those examined used a "soft-sell" approach. GCCP was compared with local consumer culture positioning (LCCP), which associates the brand with the culture of the local consumer. This strategy resonates with some consumer segments because they more readily identify with local attitudes, values, and lifestyles. Acknowledging that the 1990s represented a relatively early stage of the global diffusion of consumer culture, Alden, Steenkamp, and Batra (1999) proposed that LCCP, which reflects local values, was used more frequently than GCCP, and indeed, their findings strongly support this hypothesis.

On the basis of the above logic, it would follow that there is considerable opportunity to standardize advertising strategies and perhaps some aspects of execution. On the basis of prior evidence from Alden, Steenkamp, and Batra (1999), it follows that it may be possible to obtain a positive reaction to "soft-sell" messages across a wider range of cultures, while it is probably more difficult to realize success with standardized "hard sell" messages.

Factors Influencing GIMC

Given the theoretical bases discussed above, it seems reasonable to conceptualize GIMC as an ultimate extension of global marketing – in particular, advertising – standardization. The rationale for this is that the goal of GIMC is to create a uniform brand image in a cost-effective way that can be appealing to cross-border segments. In a strict sense, standardization may differ from the coordination of IMC in that some of the communication mix, for example, personal sales or sponsorship, cannot be standardized due to local conditions, but it is a reasonable (and feasible) starting point in a practical sense. In this light, the literature seems to suggest that there are three factors that may influence the global integration of marketing strategies of firms in multiple markets: environmental factors, strategic factors, and organizational factors. In the following sections, we discuss and extend these factors in terms of GIMC.

Environmental factors. The environmental factors include *customer similarity*, *market similarity*, *similarity of advertising infrastructure*, and the *level of competition*, all of which appear to affect the firms' willingness to implement the worldwide integration of their marketing communication mix. Collectively, we refer to these four constructs as environmental factors, as they all relate to the external environment of the cross-national markets served.

As defined here, *customer similarity* refers to the degree to which consumers in the markets served have similar lifestyles, preferences, and tastes. Cultural factors such as tastes and preferences can have a substantial impact on the viability of standardization (e.g., Jain, 1989; Laroche *et al.*, 2001; Onkvisit and Shaw, 1999). Here, a critical question relates to whether consumers exhibit a similar preference for the standardized messages regardless of culture to which they belong. It has long been argued that advertising messages should be congruent with the values of local culture (Boddewyn, Soehl, and Picard, 1986; Belk, Bryce, and Pollay, 1985; Lin, 2001; Mueller, 1987). Further, a number of empirical studies have found that advertisements reflecting (some) local cultural values are more persuasive than those that ignore them (Taylor,

Miracle, and Wilson 1997; Han and Shavitt, 1994). Thus, it is also important to examine the factors related to a culture's impact on GIMC.

Market similarity refers to the degree to which the countries served by the MNC have similar levels of economic development. In the context of GIMC, *market similarity* reflects both the extent to which managers view the individual countries as economically similar, and their perceptions of the impact of economic unification on the similarities in the overall market. As with *customer similarity*, firms are more likely to coordinate their marketing communication mix if they view the cross-national markets they serve as having similar levels of economic development (e.g., Duncan and Ramaprasad, 1995; Harvey, 1993; Samiee *et al.*, 2003).

It has been seen that the similarity of the *advertising infrastructure*, including the availability of similar media with similar costs, the presence of similar laws, and access to market research firms, also influences the level of GIMC in a market (Samiee *et al.*, 2003; Taylor and Raymond, 2000). In general, when advertising infrastructures are similar across markets, it has been found that it is more feasible for companies to engage in the global integration of their marketing communication (Jain, 1989).

The final environmental factor is the *level of competition*, which refers to the degree to which the markets served are sought after by competitors. The *level of competition* can affect the propensity to use standardized advertising messages (Samiee *et al.*, 2003; Yip, 1995). In settings where competitive conditions are intense, the coordination and integration of the various components of the marketing communication mix would make sense because it would build a consistent image that would contribute to gaining a competitive advantage, or to responding to a competitor that has built a consistent image. It should be noted that the level of competition, as defined here, does not refer to the firm's relative competitive position in each market. It is, rather, the overall degree of competition that the company faces in those markets in which it operates.

Strategic factors. The literature of marketing standardization has suggested that, in addition to environmental factors, various internal strategic factors influence the level of standardization at which a firm engages (e.g., Cavusgil and Zou, 1994; Duncan and Ramaprasad, 1995). By the same token, we propose three strategic factors, namely *global strategic orientation*, *perceived cost savings*, and *cross-border segmentation*.

The idea of a firm having a *global strategic orientation* has received considerable attention in the marketing strategy literature. In particular, Zou and Cavusgil's (2002) GMS perspective suggests that standardization of advertising is a key characteristic of firms that have a global orientation. As defined by Zou and Cavusgil (2002, p. 46), global orientation refers to "the organization-wide emphasis on success on a worldwide basis rather than on a country-by-country basis." Global orientation is part of a firm's corporate culture, and it can be viewed as a key organizational resource because it is a potential source of competitive advantage (Levitt, 1983; Ohmae, 1989; Zou and Cavusgil, 1996). Recent research evidence suggests that global strategies have a higher potential for effectiveness than they had previously. For example, Zou and Cavusgil (2002) found that firms employing a GMS achieve, on average, higher strategic and financial performance than those that do not. Moreover, Steenkamp, Batra, and Alden (2003) found that the perceived "globalness" of a brand has a positive impact on consumer perception of the brand's quality and prestige.

The second strategic factor in our model, *perceived cost savings*, refers to the firm's belief that GIMC will lead to cost savings in the production and execution of marketing communication mix programs. If the firm has a desire to save costs through GIMC, it follows that it will engage in a relatively high degree of integration. In the marketing literature, numerous authors have argued that standardization is associated with substantial cost savings. Duncan and Ramaprasad (1995) found that most of the managers surveyed did not believe that cost savings alone were a sufficient reason to justify standardized campaigns, but it is quite likely that firms that are confident that they can develop effective advertising across markets view cost savings as a key advantage. The same logic could apply for the adoption of GIMC.

Another concept related to Zou and Cavusgil's (2002) GMS perspective is the firm's belief that *cross-border segmentation* can be effective. *Cross-border segmentation* is defined as *the extent to which firms believe they can and should target consumers with similar characteristics across markets*. Several researchers have suggested that MNCs should engage in *cross-market segmentation* (Hassan and Katsanis, 1994; Miller, 1998; Shermach, 1995). A study by ter Hofstede, Steenkamp, and Wedel (1999) also suggested that cross-market segmentation could be successfully applied to the market for yogurt in Europe. To the extent that marketers engage in targeting groups on criteria that cut across borders (e.g., age, lifestyle), we predict that they will be more likely to integrate IMC globally in order to reach their customers.

Organizational factors. One of the organizational factors that has received widespread attention is the *level of control* that the parent firm has over subsidiaries, and its impact on their willingness to adopt headquarters-led strategic orientation (Samiee *et al.*, 2003). For example, it is widely believed that when control is centralized at the MNC's headquarters, more common marketing strategies will be used (e.g., Duncan and Ramaprasad, 1995; Laroche *et al.*, 2001). The basic idea is that centralized decisions help to ensure that overall corporate goals are met. If these goals involve GIMC, it is much easier to implement them centrally than it would be if local subsidiaries were granted autonomy. For example, an empirical survey by Duncan and Ramaprasad (1995) confirmed that firms with a higher degree of control emanating from headquarters were more prone to engage in standardized advertising. Similarly, Laroche *et al.* (2001) found that those MNCs with highly centralized control engaged in more standardization.

The *size of the firm* is defined here as *the subsidiary's annual sales volume*. As Samiee *et al.* (2003) noted, prior research on international business suggests that larger firms are more likely to expand into foreign markets and to have a standardized worldwide image. Additionally, work on entry to foreign markets suggests that large firms prefer wholly owned subsidiaries, partly to ensure that their general corporate strategy is carried out uniformly (Erramilli and Rao, 1993). At the subsidiary level, it can be argued that it is likely that larger firms will also be more prone to engage in closely coordinated activities

Several researchers have found that *level of international experience* is also helpful to firms in planning an international strategy. Cavusgil and Zou (1994) found empirical support for the idea that more experienced exporters have higher levels of success in international markets. Additionally, Zou and Cavusgil (2002) found a significant link between the firm's *level of international experience* and its use of a GMS, including a global promotional strategy.

Conclusion

The increasing availability of global mega-agencies enables firms to adopt the global integration of diverse marketing communication mixes across border and facilitates the development of globally integrated marketing communication or GIMC. GIMC is defined as *an ultimate extension of global marketing standardization*. In doing so, two theoretical frameworks are explained: GMS theory and global brand positioning theory. Wherever possible, more and more firms are interested in coordinating their marketing communication mix in multiple markets in order to create a uniform and strong brand image and ultimately to increase brand equity. Also, such global integration decreases costs in developing and maintaining marketing strategies and programs, and eases the execution of common advertizing campaigns to cross-border segments that share similar preferences.

See also *base of the pyramid markets: culture insights and marketing implications; international advertising – is there still a standardization versus local adaptation debate?; global branding: three keys for global brand success; society, culture, and global consumer culture; standardization/adaptation of international marketing strategy*

Bibliography

Agarwal, M. (1995) Review of a 40-year debate in international advertising. *International Marketing Review*, **12** (1), 26–48.

Alden, D.L., Steenkamp, J.-B.E.M., and Batra, R. (1999) Brand positioning through advertising in Asia, North America, and Europe: the role of global consumer culture. *Journal of Marketing*, **63** (1), 75–87.

Belk, R.W., Bryce, W.J., and Pollay, R.W. (1985) Materialism and status appeals in Japanese and U.S. print advertising. *International Marketing Review*, **2** (12), 38–47.

Boddewyn, J.J., Soehl, R., and Picard, J. (1986) Standardization in international marketing: is ted levitt in fact right? *Business Horizons*, **29** (6), 69–75.

Cavusgil, S.T. and Zou, S. (1994) Marketing strategy-performance relationship: an investigation of the empirical link in export market ventures. *Journal of Marketing*, **58** (1), 1–21.

Duncan, T. and Ramaprasad, J. (1995) Standardized multinational advertising: the influencing factors. *Journal of Advertising*, **24** (3), 55–68.

Elinder, E. (1961) How international can European advertising be? *Journal of Marketing*, **29** (2), 7–11.

Erramilli, M.K. and Rao, C.P. (1993) Service Firms' international entry-mode choice: a modified transaction-cost analysis approach. *Journal of Marketing*, **57** (3), 19–38.

Goldsmith, M. (2009) Marketing in the Age of Turbulence, BusinessWeek (June 5), available at: http://www. businessweek. com/ managing/ content/ jun2009/ca2009065_435823.htm.

Gould, S.J., Lerman, D.B., and Grein, A.F. (1999) Agency perceptions and practices on GIMC. *Journal of Advertising Research*, **39** (1), 7–20.

Han, S.-P. and Shavitt, S. (1994) Persuasion and culture: advertising appeals in individualistic and collectivistic societies. *Journal of Experimental Social Psychology*, **30** (4), 326–350.

Harris, G. (1994) International advertising standardization: what do the multinationals actually standardize? *Journal of International Marketing*, **2** (4), 13–30.

Harvey, M.G. (1993) A model to determine standardization of the advertising process in international markets. *Journal of Advertising Research*, **33** (4), 57–64.

Hassan, S.S. and Katsanis, L. (1994) Global market segmentation strategies and trends, in *Globalization of Consumer Markets: Structures and Strategies* (eds S.H. Salah and E. Kaynak), International Business Press, New York, pp. 47–62.

ter Hofstede, F., Steenkamp, J.-B.E.M., and Wedel, M. (1999) International market segmentation based on consumer-product relations. *Journal of Marketing Research*, **36** (1), 1–17.

Jain, S.C. (1989) Standardization of international marketing strategy: some research hypotheses. *Journal of Marketing*, **53** (1), 70–79.

Kanso, A. and Nelson, R.A. (2002) Advertising localization: overshadows standardization. *Journal of Advertising Research*, **42** (1), 79–89.

Kitchen, P.J. and Schultz, D.E. (1999) A multi-country comparison of the drive for IMC. *Journal of Advertising Research*, **39** (1), 21–38.

Laroche, M., Kirpalani, V.H., Pons, F., and Zhou, L. (2001) A model of advertising standardization in multinational corporations. *Journal of International Business Studies*, **32** (2), 249–266.

Levitt, T. (1983) The globalization of markets. *Harvard Business Review*, **61** (3), 92–102.

Lin, C.A. (2001) Cultural values reflected in chinese and american television advertising. *Journal of Advertising*, **30** (4), 83–95.

Melewar, T.C. and Saunders, J. (1999) International corporate visual identity: standardization or localization? *Journal of International Business Studies*, **30** (3), 583–597.

Miller, T. (1998) Global segments from 'Strivers' to 'Creatives'. *Marketing News*, **32** (5), 11.

Moriarty, S.E. and Duncan, T.R. (1990) Global advertising: issues and practices. *Current Issues and Research in Advertising*, **13**, 313–341.

Mueller, B. (1987) Reflections of culture: an analysis of Japanese and American advertising appeals. *Journal of Advertising Research*, **27** (3), 51–59.

Ohmae, K. (1989) Managing in a borderless world. *Harvard Business Review*, **67** (3), 152–161.

Okazaki, S. (2005) Searching the web for global brands: how American brands standardise their websites in Europe. *European Journal of Marketing*, **39** (1/2), 87–109.

Okazaki, S., Taylor, C.R., and Zou, S. (2006) Advertising standardization's positive impact on the bottom line: a model of when and how standardization improves financial and strategic performance. *Journal of Advertising*, **35** (3), 17–33.

Onkvisit, S. and Shaw J.J. (1987) Standardized international advertising: a review and critical evaluation of the theoretical and empirical evidence. *Columbia Journal of World Business*, **22**, 43–54.

Onkvisit, S. and Shaw, J.J. (1999) Standardized international advertising: some research issues and implications. *Journal of Advertising Research*, **39** (6), 19–24.

Samiee, S., Jeong, I., Pae, J.H., and Tai, S. (2003) Advertising standardization in multinational corporations: the subsidiary perspective. *Journal of Business Research*, **56** (8), 613–626.

Shermach, K. (1995) Portrait of the world. *Marketing News*, **29** (18), 20–21.

Steenkamp, J.-B.E.M., Batra, R., and Alden, D.L. (2003) How perceived brand globalness creates brand value. *Journal of International Business Studies*, **34** (1), 53–65.

Synodinos, N.E., Keown, C.F., and Jacobs, L.W. (1989) Transnational advertising practices: a survey of leading brand advertisers in 15 countries. *Journal of advertising research*, **29** (2), 43–50.

Szymanski, D.M., Bharadwaj, S.G., and Varadarajan, P.R. (1993) Standardization versus adaptation of international marketing strategy: an empirical investigation. *Journal of Marketing*, **57** (4), 1–17.

Taylor, C.R., Miracle, G.E., and Wilson, R.D. (1997) The impact of information level on the effectiveness of U.S. and Korean television communication. *Journal of Advertising*, **26** (1), 1–18.

Taylor, C.R. and Raymond, M.A. (2000) An analysis of product category restrictions in advertising in four major East Asian markets. *International Marketing Review*, **17** (3), 287–304.

Tharp, M. and Jeong, J. (2001) Executive insights: the global network communications agency. *Journal of International Marketing*, **9** (4), 111–131.

Yip, G.S. (1995) *Total global strategy: managing for worldwide competitive advantage*, Prentice Hall, Englewood Cliffs, NJ.

Zou, S. and Cavusgil, S.T. (1996) Global strategy:a REVIEW and an integrated conceptual framework. *European Journal of Marketing*, **30** (1), 52–70.

Zou, S. and Cavusgil, S.T. (2002) The GMS: a broad conceptualization of global marketing strategy and its effect on firm performance. *Journal of Marketing*, **66** (4), 40–56.

database marketing

John A. McCarty

Introduction

Database marketing involves the analysis of customer transaction data, along with other customer information (e.g., demographics), to segment customers (*see* MARKET SEGMENTATION AND TARGETING), and to develop marketing strategies (*see* MARKETING STRATEGY) to some or all of those segments. A key aspect of this description is that database marketing almost always involves the analysis of transaction data, either alone or with other kinds of customer data. Transaction data include information about the actual purchases made by consumers, such as when a purchase was made, the assortment of items that were purchased, and the total amount that was spent during a particular purchase. Thus, database marketing is generally employed by marketers having access to such information about individual customers, including catalog marketers, retailers, credit card companies, and others that engage in direct contact of one sort or another with customers. Nonprofits (e.g., charities, associations) have also utilized database marketing. A charity, for example, has information about donors, including how much they have donated over the years, the amount of their largest donation, and the number of times they have donated.

Among marketing firms, direct marketers (e.g., L.L. Bean) were the first to employ database marketing techniques. Their earliest efforts were largely informal in that direct marketers tended to note the patterns and preferences of their customers and to utilize knowledge of these in their marketing to repeat customers (Baier, Ruf, and Chakraborty, 2002). As the number of customers for these firms grew to a level that would make such informal methods impractical, computers and analytical procedures were developing that could automate the tracking of customer behavior.

Today, database marketing involves a number of analytical procedures, both basic and sophisticated, that enable direct marketers to track the behavior patterns of their customers. There has also been an increase in the use of database marketing by firms that are not traditional direct marketers, such as credit card providers, automobile rental agencies, brick and mortar retailers, casinos, and hotel chains. Virtually any firm that routinely captures transaction information as part of the way it does business can engage in database marketing.

Database Marketing and Traditional Mass Marketing

Firms that engage in database marketing differ from most traditional mass marketers that produce packaged goods (e.g., ketchup, paper towels) in two important and related ways. First, database marketers have individual customer-level information. Since they generally deal with their customers in a direct manner, they can identify their customers individually, including customers' names and addresses. Second, database marketers have the information about individual customer's transactions with the firm. That is, every time a specific customer makes a purchase, the marketer collects information about how many different items were purchased, how many of each of these items were purchased, how the customer made the payment (i.e., credit or check), and so on. This individual level information allows database marketers to serve very small segments of their customers and to personalize and customize their marketing efforts to individuals. Furthermore, given that they have the actual transaction data of customers, database marketers can segment and develop strategies based on the actual behavior of customers rather than on characteristics such as demographics.

What database marketers can do with these data contrasts with the abilities of traditional mass marketers. A typical mass marketer, like the ones that produce packaged goods, sells to consumers through intermediaries. Mass marketers do not know the names or characteristics of their individual customers and have no information about the specific purchases, as these transactions occur in supermarkets, drugstores, and/or mass merchandisers. Mass marketers can make inferences about the characteristics of their customers based on survey research and other marketing research studies. Although such studies can provide

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

these marketers with insights about their typical consumer, they cannot tell them about the behaviors of individual consumers. Survey research may inform a mass marketer that the consumers who purchase its brand are generally of a particular age group and gender. Other types of marketing research may tell the mass marketer such things as on what occasions consumers typically use the product, in what quantities they use it, or even how they feel about it (e.g., the pros and cons of its brand relative to other brands). Although this information is extremely valuable to mass marketers in the development of their marketing strategies to the broad segments that they serve, it cannot provide them with ways of marketing to specific customers.

Although it may seem that database marketers have enormous advantages over mass marketers, the ability to market to individual customers comes with certain costs. Developing and maintaining a customer database can be a costly activity. Designing and implementing marketing strategies to small segments of customers and developing customized promotional materials can be very expensive. Therefore, the revenues that firms employing database marketing techniques receive must cover the costs of these efforts. For the vast majority of traditional marketers, the selling of their products through mass marketing outlets such as supermarkets makes sense, given their profit margins. Moreover, it is convenient for most consumers to buy their packaged goods in one place (e.g., local supermarkets); therefore, consumers are generally not interested in purchasing such items as toothpaste, canned beets, and toilet paper in a direct way from the different marketers of these goods. Therefore, for both the firms producing most packaged goods and the consumers of those goods, the way these things are bought and sold makes reasonable sense.

Considering marketing activities at a broad level, it can be said that database marketers and traditional mass marketers engage in similar activities; that is, both types of marketers engage in segmentation and the development of marketing strategies to serve these segments. The crucial difference is the kind of information that is used to engage in these marketing activities. Database marketers have information that allows them to segment on the specific behaviors of their customers and to develop customized strategies to narrow segments or even to individuals. Traditional mass marketers have information from marketing research that allows them to segment on variables that relate to consumption (e.g., demographic information) and develop strategies based on the inferred behavior of large segments of consumers.

Questions that Database Marketing may Address

As noted, database marketing involves the quantitative analysis of customer transaction information in an effort to engage in marketing strategy and tactics with respect to those customers. Three examples aid in understanding the range of questions that database marketing may address.

A common issue for a database marketer is how to segment its database such that it can identify the very best customers, those that it may be able to covert to this top group of customers, and even those toward whom it may want to discontinue marketing efforts. (Hughes, 2005). The very best customers (gold customers, according to Hughes terminology) are those who are doing a lot of business with the company and these are the ones on which the marketer may want to spend more customer service dollars. Those who are just below gold customers are the ones that the company may want to market to heavily, with an eye toward increasing their business to the level of gold customers. At the other extreme, there are customers who do very little business with the marketer. The revenues obtained from these customers may be so little as to not really justify any marketing effort by the firm. Analysis of the revenues obtained from different customers in the data file and their patterns of transactions can help database marketers place their customers in these various segments.

A very common question for catalog marketers is, "who among our total customers should receive an upcoming mailing?" This sort of question is a very serious one for marketers relying on the mail to conduct business in that the postage rate generally makes mailing to the total customer file cost prohibitive. Even if the cost is much less than a dollar per piece, mailing

to a file of tens of thousands of customers is an expensive proposition when the typical response rate is less than 5% for a single mailing. This problem is a segmentation and targeting issue in that a marketer would like to mail to the people in the database who are the most likely to respond to the mailing. Identifying those most likely to respond to a particular mailing from transaction variables in the customer data file is a typical database marketing problem. Analysis of the customers who have responded to similar mailings in the past compared with those who did not is one way of evaluating who should receive an upcoming mailing.

A different sort of database marketing effort is in the area of cross-selling. Cross-selling involves efforts toward marketing one product or service to customers currently purchasing other products or services from a marketer. The nature of the information in a customer data file allows marketers to evaluate cross-selling opportunities. Marketers can evaluate the pattern of purchases for the typical customer and determine if there are different products that customers tend to consistently buy together. A supplier of clothing, for example, may note that customers who buy dress shirts from it also tend to buy ties. A cross-selling opportunity would be to identify those customers who have been buying shirts, but not ties, and attempt to market ties to these customers through a promotional effort.

Database Marketing and Data Mining

Database marketing is the application of data mining techniques in the context of marketing strategy and tactics. According to Berry and Linoff (2000), "data mining is the process of exploration and analysis, by automatic or semi-automatic means, of large quantities of data in order to discover meaningful patterns and rules (p. 7)". Data mining, therefore, can be thought of as an approach to understanding and interpreting data; when marketers engage in mining their customer data, it is database marketing. However, data mining has many applications – not only marketing or business applications but also in such diverse areas as health care, engineering, the military, and public policy, to name a few. Data mining is applicable to any situation where there is an enormous amount of quantitative information and there is belief that understanding the patterns of the data will provide useful insights.

Data mining techniques tend to be data driven rather than hypothesis driven (*see* HYPOTHESIS TESTING RELATED TO DIFFERENCES – PARAMETRIC TESTS). In other words, the typical data mining analytical procedure searches for patterns and relationships in the data. In contrast, hypothesis-driven research begins with beliefs about the patterns that may exist in the data. The data are used to test the viability of a hypothesis. The advantage of the data-driven approach is that one need not have a priori assumptions about the relationships; one is searching for any patterns that may exist. Therefore, one may discover relationships that were not anticipated and these serendipitous findings may prove to be meaningful and useful to the database marketer. The very real danger with such data-driven approaches is that one may be capitalizing on chance relationships in the data. Without a hypothesis based on sound theory, it is far more likely that one will find relationships because of the search for any relationship; some of the found relationships may exist because of chance occurrences rather than meaningful and robust relationships that will maintain across time. It is therefore critical that database marketers understand this important aspect of their analysis approach and take steps to evaluate the reliability of their findings.

Database Marketing and Relationship Marketing

Database marketing is closely associated with relationship marketing (*see* CUSTOMER RELATIONSHIP MANAGEMENT). Although definitions of relationship marketing vary from very narrow views to broad conceptualizations (Parvatiyar and Sheth, 2000), a primary focus of relationship marketing is on the long-term interactions between buyers and sellers rather than on the individual transaction event. So, for example, an automobile dealership approaching its business in a relationship manner will think in terms of servicing its customers' transportation needs over years rather than the sale of a car on a specific day. This would entail a consideration of postsale services and attention to the customers

such that they will buy future cars from the dealer. Those promoting relationship marketing argue that such long-term focus will be beneficial to both the buyer and seller.

Direct marketers were some of the earliest marketers to think in terms of long-term relationships with their customers. To a great extent, direct marketers' focus on the long term was out of necessity. For a typical mailing, the postal costs are very high relative to the revenues resulting from that specific mailing. This cost-to-revenue relationship dictates that direct marketers think in terms of the long-term profitability of the customers acquired from the mailing in order to justify the mailing. An illustration will help make this point. Imagine that a marketer is going to do a mailing to 50 000 people. The cost of the mailing is $0.75 per piece; the total mailing cost is therefore $37 500. The response rate is 2.5% (a typical response rate for a cold mailing), which amounts to 1250 people responding. The average revenue per response (with all costs except for mailing subtracted) is $30. So, the total revenue before the mailing costs are subtracted is $37 500. Thus, it appears that the marketer would make no money from the mailing, to the extent that it thinks only in terms of this mailing and the revenues received from this mailing. If the direct marketer thinks about the long term, it will consider how much revenue that the 1250 responders will generate over time. This revenue will not only be in terms of those responders' future purchases but also the revenues generated by people they may refer to the marketer. These 1250 people may generate a lot of profits for the company over the next several years, to the extent that the direct marketer practices sound relationship marketing. This way of thinking relates to the concept of the lifetime value of the customer (*see* CUSTOMER LIFETIME VALUE (CLV)). The lifetime value of a customer is the net present value of the revenues that will be received from a customer over a period of time in the future (usually 3–5 years). In this example, the direct marketer may have only received $30 in revenues from the average customer for this mailing, but its lifetime value analysis may estimate that it will receive hundreds of dollars from the average responder when considered across time.

Even if a database marketer does not have postal costs, there are very sound reasons for taking the long view with customers. Maintaining a database and engaging in the customized marketing efforts associated with database marketing is expensive. Such expenses will generally only pay for themselves across time. If there is a great deal of churn among a marketer's customer base, the database marketer's analytical procedures may be wasted as the customer database at any given time may contain mostly newly acquired customers.

As a general rule, database marketing works best when the majority of people in the database desire a relationship with the marketer. Alternatively, if a significant number of customers are price-oriented customers, the marketer will likely lose them as soon as another firm offers a lower price. Therefore, engaging in database marketing efforts like loyalty programs when customers are price sensitive may likely not be a worthwhile effort.

The Analytical Procedures of Database Marketing

There are a number of commonly used analytical procedures in database marketing. Although these procedures may differ on the specifics, in general, they are designed to uncover relationships among the various transaction variables in the database. Among other purposes, the analytical procedures can be used to segment customers into groups based on purchase patterns (e.g., frequently purchasing customers, the biggest spenders, and the most loyal customers); identify the customers that are most likely to respond to a specific offer; or identify the best prospects for a cross-selling effort. The analytical procedures of database marketing vary from fairly simple methods that have been around for a long time, such as recency–frequency–monetary (RFM) analysis, to very sophisticated methods that have recently emerged, such as neural network models.

RFM analysis. Years before database marketing was conceptualized in any sort of formal way, direct marketers (e.g., catalog companies) had observed that customers who had recently purchased from them, those who had frequently

purchased from them, and those who had spent a fair amount of money with them were the best prospects for new offers (Baier, Ruf, and Chakraborty, 2002). Over time, these informal observations grew into the use of the database marketing analytical procedure known as *RFM analysis*. RFM analysis involves the consideration of how recently customers have purchased from the marketer, measured in number of days, weeks, or months since last purchase (recency); how frequently customers have purchased from the marketer, measured as how many times they have purchased from the marketer (frequency); and the total amount of money they have spent with the marketer (monetary value). Generally, these three variables are used to predict which customers in a data file are the most likely to respond to a particular offer.

There are a variety of approaches used in RFM analysis, differing in terms of how the relative weights of the RFM variables are determined. One common approach to determining this weighting is known as *hard coding* (Drozdenko and Drake, 2002). Hard coding involves the development of a simple linear equation of the RFM variables, where weights for these three variables are assigned as a matter of judgment and past experience. For example, a marketer may reason from past experience with similar offers that recency should be assigned a weight of 2, frequency should be assigned a weight of 1, and monetary value should be weighted 0.5. Therefore, the linear equation provides a weighted score for each person in the data file; those with the higher scores should be the ones who are more likely to respond to a particular offer. The marketer can decide how far down in the data file it should go, for example, mail to 30% of the file, 50% of the file, and so on. Again, past experience would guide the marketer in this decision. Presumably, a greater profit would be realized from mailing to a smaller group of higher probability customers, based on RFM, than mailing to everyone in the data file. The relative weights of the RFM variables could also be determined empirically from the results of a similar mailing in the past. Regression analysis could be used to determine these weights.

An alternative approach to RFM analysis is a method advocated by Hughes (2005). One key feature of Hughes' approach involves the use of a test mailing to a sample of customers in the data file prior to the selection of members of the data file who will receive the mailing. Hughes' approach begins with the sorting of the entire customer data file according to how recently the customers have purchased from the company. The customers are then divided into quintiles according to recency with these quintiles assigned a number from 5 to 1. Thus, 20% of the customers who most recently purchased are assigned a recency score of 5, the next quintile of customers are assigned a recency score of 4, and so on. Next, within each of the recency quintiles, the customers are sorted according to how frequently they have purchased from the marketer, placed in quintiles according to that variable, and given a score of 5 to 1. Finally, each of the frequency quintiles within the recency quintiles are sorted by total amount of money spent with the marketer, placed in monetary value quintiles, and assigned scores of 5 to 1. This procedure results in 125 unique cells with RFM scores of 555 (the most recent, most frequent within recency, and the most money spent within the other two variables) down to 111 (the least recent, the least frequent, and the least money spent).

Hughes suggests that a test mailing be sent to a sample of approximately 10% from each of the RFM cells. An analysis of the revenues and the costs of mailings will provide the response percent needed in order to break even. The response percent in the test mailing for each of the 125 RFM cells is compared to the break-even percent. The marketer then mails to the customers who were not in the test mailing, but only those from the RFM cells that broke even in the test. Typically, the number of cells that are mailed to is far less than the total of 125 cells; however, the profit is higher because of the mailing costs that are saved by not mailing to the unproductive cells.

Although more sophisticated analytical procedures have been developed in recent years, RFM analysis continues to be popular for a number of reasons. Ease of use and simplicity are often cited as reasons for its popularity, as well as the fact that it is a method that decision makers can easily understand (McCarty and Hastak, 2007). The simplicity of RFM is, however, often why

it is criticized as a database marketing technique. Critics point out that there are really no sound reasons for limiting the analysis to these three variables when there might be other transaction variables in the customer file that would help identify likely responders.

CHAID analysis. An analytical procedure that has been used by database marketers for segmentation and targeting purposes is chi-square automatic interaction detector (CHAID). Although this analytical procedure was developed for purposes unrelated to segmentation, it has proved to be very useful to database marketers.

CHAID is a classification method that can also be referred to as *tree analysis*. It is often used to find the most likely responders to target for a mailing or other marketing effort. For the purposes of this explanation, let us say that a direct marketer wants to do a major mailing and decides to conduct a test mailing of about 10% of its entire house data file of 1 000 000; the results of the test mailing will be used to identify those in the rest of the data file to receive the mailing. The response results creates the criterion variable to use for the CHAID analysis; this variable is the dichotomous variable of those among the 10% who responded to a test mailing and those who did not respond. In this example, let us say that the overall response rate was 5% of those who were sent the test mailing (i.e., 5000 responders). A number of independent variables are then tested to see which of them are most useful in discriminating between responders and nonresponders. These independent variables can be any variables that the marketer might believe are related to response likelihood, such as the RFM variables, and other variables as well, such as method of payment (i.e., credit card or check), number of different items purchased from the marketer, and so on.

The CHAID procedure begins with a node of all the customers in the test mailing. CHAID searches among the identified independent variables for the one that best discriminates among the customers with respect to the criterion variable (i.e., respond or did not respond to the mailing). For the purposes of this illustration, let us say that the best discriminating independent variable is how frequently customers have purchased from the marketer. The procedure will split the original set of people who have

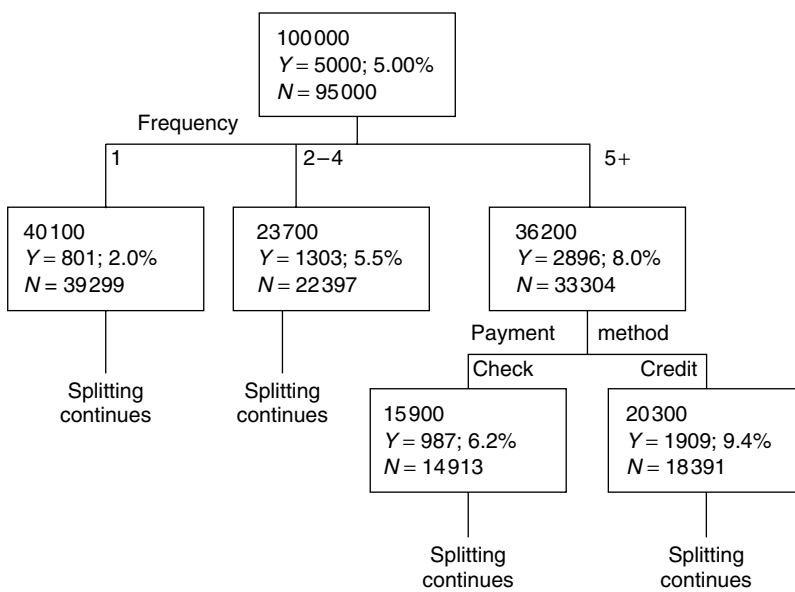


Figure 1 Example of a CHAID classification tree. This is the initial splitting of the CHAID analysis; splitting may continue until no significant differences are found.

sent the test mailing into two or more groups, depending on how this independent variable discriminates with respect to the criterion variable. For example, it may group those who have purchased five or more times together into one node, those who have purchased two to four times in a second node, and those who have purchased only once in a third node, as shown in Figure 1. Although the overall response rate was 5%, the response rate for these three nodes will be different on the criterion variable. Perhaps it is 8% for those who had purchased five or more times, 5.5% for those who had purchased two to four times, and 2% for those who had only purchased one time.

The CHAID procedure continues splitting the groups using this same logic. For each of these three nodes identified by the first independent variable (i.e., frequency), CHAID identifies the variable that best discriminates among the people in each of those groups. So, for example, among those who had purchased five or more times, the best discriminating variable might be whether they purchased with a credit card or check; the procedure will divide this group by the method of payment variable and the two resulting groups will have a different level on the criterion variable. The procedure may divide those who purchased two to four times on a different independent variable, as CHAID considers each of the groups from the original split independently in its search for the next best discriminating variable. This splitting of nodes continues until there are no more splits that would be statistically significant.

The CHAID procedure begins with a node of everyone who received the test mailing and ends with a number of nodes or groups of similar individuals with respect to the various independent variables. In this example, let us say that there are 30 final nodes. The nodes will differ with respect to their response rate. If the independent variables are indeed useful in discriminating between responders and nonresponders, then the response rate of the 30 different nodes may vary from percentages well above the response rate of the total group to some that are well below that of the total. The marketer can then place the rest of the total data file (i.e., the 90% who did not receive the test mailing) into groups or nodes according to their values on the independent variables that defined the 30 nodes. At this point, the marketer can decide which of the nodes should be sent the mailing. This decision may be a function of how many people they can mail to given a certain budget or according to a break-even analysis. In other words, a response percentage that would be necessary to break even can be determined by the revenue and costs of mailing; the marketer can then mail to only those nodes that had that percentage in the test mailing.

CHAID is similar to Hughes' (2005) approach to RFM analysis in that it places customers into cells or nodes; however, CHAID determines these nodes as a function of how the independent variables behave empirically with respect to the criterion variable. Moreover, as noted, CHAID is not restricted the RFM variables as independent variables. CHAID has been shown to be superior to hard coding RFM (Levin and Zavari, 2001; Magidson, 1988); it has been shown to perform better than Hughes' approach to RFM when the overall response rate is low, but similar to RFM when the response rate is relatively high (McCarty and Hastak, 2007).

Multiple linear regression and logistic regression. Multiple linear regression (*see* MULTIPLE REGRESSION) is a general statistical procedure that is used to predict a continuous dependent variable from a set of independent variables. Multiple regression can be useful to database marketers in the instances when they are interested in understanding the relationship among a set of transaction variables in their data file and the criterion variable of interest is not a discrete variable (e.g., response/no response to a mailing).

An example of when multiple regression may be used is when a database marketer would like to know what predicts the dollar amount of business that customers will spend with them. For example, a casino may want to know what variables are related to higher spending among their customers. The reward cards given to customers provide estimates of this information. The criterion variable is amount of money gambled (often estimated for table games; accurately provided for slot machines if the customers insert their

card in the machines). There are a variety of independent variables that could be used as predictors including length of stay in the casino hotel, games that a customer plays (e.g., slot machines, table games), number of customer visits in a year, distance from customer's hometown to the casino, and demographic characteristics of a customer's hometown (available from Census Bureau data). The casino database marketer can use these independent variables to predict the criterion variable of money spent; the weights for the independent variables provide insights about the spending habits of its best customers. For example, the marketer may discover that the number of visits a customer makes is not as important as length of stay of visits in predicting spending level, or that bigger spenders play certain table games, and so on.

Another benefit from this sort of regression analysis is that it can provide information about customers who are good prospects for additional marketing. Customers who are predicted to be spending more than they actually are, given their levels on the important independent variables, can be sent promotional materials in an effort to increase their business with the casino.

Logistic regression (*see* LOGIT MODEL) is also utilized by database marketers. Logistic regression models a dichotomous dependent variable with a set of independent variables. Thus, it is useful for database marketing in instances when a marketer would like to predict response/no response to a marketing effort. A particularly useful aspect of logistic regression is that although the criterion variable is dichotomous (i.e., either a customer responded or did not); the resultant predicted values are probabilities ranging between 0 and 1. Thus, if logistic regression is performed on a test mailing or the results of a previous mailing, the analysis provides a weighted function of independent variables that best predict the response variable. The weights of the independent variables provide the information to develop the response probabilities for other customers (e.g., those who were not sent a test mailing). The marketer can decide what level of predicted response probability to use in deciding who should receive the mailing. A break-even analysis may tell the marketer that it will be profitable to mail to anyone in the data file who has a predicted response probability of 30% or higher, for example.

Logistic regression is similar to RFM and CHAID in that it is typically used to determine who should receive a mailing or other marketing effort, based on information from a previous mailing or a test mailing. It differs from the RFM and CHAID in that it does not group people into cells or nodes, rather it provides an individual score (i.e., response probability) for each person.

Neural networks. Neural network analysis is one of the most recent advances in database marketing analytical techniques. These procedures have been widely used in a variety of data mining applications in business and nonbusiness areas (Berry and Linoff, 1997). Neural networks find patterns in data using an approach that mimics the human brain in how it makes connections. The procedure uses an iterative process to "learn" from training data sets; it develops a sense of these patterns and sharpens them during this learning process, analogous to how a human may learn and understand across time. Neural networks have advantages over many other analytical techniques used in database marketing in that they do not follow a specific empirical model (e.g., the linear model) and can be adapted to a wide variety of problems. A main disadvantage of neural networks is that the results of the analyses are not always easily understood (Berry and Linoff, 1997); therefore, they are not always easy to explain to decision makers.

Prospecting for New Customers

Most database marketing activities are in support of marketing efforts directed toward a company's current customers. Companies do need to acquire new customers, however, in that there will always be some attrition of customers, no matter how good are a firm's relationship marketing practices. A key difference between the use of database marketing for current customers and the prospecting for new customers is that a firm does not have transaction information for customers that it is hoping to acquire.

In spite of this limitation, database marketing can help in new customer acquisition and there

are a variety of approaches to this. One approach is for a company to develop a demographic profile of their best customers and then acquire lists of people who fit this profile to "cold call." The Census Bureau and commercial vendors (e.g., Nielsen Claritas) provide information about geographic areas that can help database marketers in their prospecting efforts. For example, if a resort hotel finds that its typical customers come from geographic areas with particular demographics and lifestyle characteristics, it can use such information to aggressively market in areas with similar demographic and lifestyle profiles. Therefore, the use of database marketing for prospecting is generally a matter of a firm understanding their own customers, then seeking others similar to their best customers.

Summary

Database marketing involves the application of data mining techniques to the analysis of a company's customer transaction data in an effort to segment customers and to develop marketing strategies to customers. Database marketing is closely aligned with the concept of relationship marketing in that those who engage in database marketing generally need to consider their long-term relationships with customers for the cost of developing, maintaining, and mining a database to be profitable. Furthermore, since the information available in customer databases tends to allow marketers to provide personalized marketing strategies to their customers, such customized efforts tend to foster strong relationships with customers. A number of analytical techniques are utilized in database marketing to help marketers discover meaningful patterns in their customer data; some of these techniques have been available for a long time (e.g., RFM analysis) while others have only recently been made available to the database marketing (e.g., neural network analysis). It is likely that the use of database marketing will grow as more types of organizations learn about its value in developing close relationships with customers.

Bibliography

Baier, M., Ruf, K.M., and Chakraborty, G. (2002) *Contempory Database Marketing: Concepts and Applications*, Racom Communications, Evanston.

Berry, M.J.A. and Linoff, G. (1997) *Data Mining Techniques for Marketing, Sales, and Customer Support*, John Wiley & Sons, Inc., New York.

Berry, M.J.A. and Linoff, G. (2000) *Mastering Data Mining: The Art and Science of Customer Relationship Management*, John Wiley & Sons, Inc., New York.

Drozdenko, R.G. and Drake, P.D. (2002) *Optimal Database Marketing: Strategy, Development, and Data Mining*, Sage Publications, Thousand Oaks.

Hughes, A.M. (2005) *Strategic Database Marketing*, 3rd edn, McGraw-Hill, New York.

Levin, N. and Zahavi, J. (2001) Predictive modeling using segmentation. *Journal of Interactive Marketing*, **15**, 2–22.

Magidson, J. (1988) Improved statistical techniques for response modeling: progression beyond regression. *Journal of Direct Marketing*, **15**, 494–501.

McCarty, J.A. and Hastak, M. (2007) Segmentation approaches in data-mining: a comparison of RFM, CHAID, and logistic regression. *Journal of Business Research*, **60**, 656–662.

Parvatiyar, A. and Sheth, J.N. (2000) The domain and conceptual foundations of relationship marketing, in *Handbook of Relationship Marketing* (eds J.N. Sheth and A. Parvatiyar), Sage Publications, Thousand Oaks, pp. 1–38.

the advertising budget

Douglas C. West

Introduction

The advertising budget is the financial statement and program put before top management for approval for spending on media, advertising production, and advertising services in order to meet advertising objectives. Expenditures on advertising can be enormous and often dwarf all other forms of marketing activity, and yet there is no evidence of widespread sophisticated practice (Kissan and Richardson, 2002). Many of the methods employed are historic and intuitive and have not been taught so much as observed and classified by researchers. To a great extent a muddled business reality has been formalized and compartmentalized by researchers.

Methods

The leading approaches and specific methods, with their percentage use as indicated by surveys are listed below (see Synodinos, Keown, and Jacobs 1989; Prendergast, West, and Shi 2006; Farris and West, 2007; Belch and Belch, 2009, pp. 226–42):

Judgmental methods (ca 25–30% of budgets):

- Arbitrary: The advertising budget is solely determined on the basis of what is "felt" to be necessary.
- Affordable: The organization determines what it can spend on areas such as production and operations and then decides how much it can then afford for advertising.

Objective and task (ca 25–30% of budgets):

- Spending is in accordance with what is required to meet the advertising objective(s) which are ranked by importance. Tasks are agreed upon to meet these objectives and then costs estimated. If the campaign cannot be afforded, objective(s) that are of lower importance are eliminated until the budget can be afforded.

Measurement (ca 20% of budgets):

- ROI: Advertising is considered an investment, and money is spent to the point where the ROI is diminishing.
- Incremental: The budget is allocated in a series of tests (e.g. spend 5% more or 5% less) mainly using direct and interactive media such as direct mail, the Internet, and tools like barcode data. Spending is increased or decreased in line with the results achieved.
- Quantitative models: Computer simulation models involving statistical techniques such as multiple regression analysis are used.

Percentage of sales (ca 15% of budgets):

- Percentage of last year's sales: Set as a percentage of previous financial year's sales.
- Percentage of anticipated sales: Set as a percentage of the firm's anticipated sales.
- Unit sales: The organization allocates a fixed percentage of unit price for advertising and then multiplies this amount by projected sales volume (e.g. 5% unit price × 200,000 units forecast).

Competitive (ca 5–7% of budgets):

- Competitive absolute: The budget is set in line with that of the closest rival.
- Competitive relative: All the competitors in the market tend to spend in line with their market share.

Choices

From an academic perspective, advertising budgeting studies can be largely compartmentalized as either "method" or "organizational" in their focus. Method studies have examined the effects of size, market type, performance, and country on advertising budgeting, but no definitive conclusions have arisen. Studies have given a broad measure of support for the increasing use of more complex methods namely, the "objective and task," and to a lesser extent, measurement methods, from the 1980s onward. There are some signs that Business-to-Consumer (B2C) advertisers, larger

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

companies, better performers, and those within more highly developed economies are more likely to use measurement methods, but the evidence is not strong. Organizational studies have examined the political processes and structures that affect setting any budget.

When it comes to managers, the advertising budgeting decision is part of their overall decision-making and *satisficing*, which involves seeking alternatives and choosing the first one that exceeds your expectations, is a common approach. Managers are often constrained by a multitude of factors, but particularly by lack of time, data, and resources and the politics of their own organization. They know how to set their advertising budget ideally, but yet have to compromise in order to develop the budget in practice. So while they might preferably want to do some sophisticated econometric modeling, instead they stick to the well tried and trust percentage of sales method for example, because it produces a budget good enough to fit their needs and saves a lot of time.

Summary

The advertising budget is an important decision. Spend too much and resources will be wasted. Spend too little and opportunities will be lost. The main methods are judgmental, objective and task, measurement, percentage of sales, and competitive and they can be used singularly or in combination. Academics have largely investigated the effect of market and firm variables on choices and issues related to organizational structure and politics. Managers often make their choices of budgetary method based upon satisficing. Faced with scarcities of all kinds (information and time in particular), managers regularly compromise the ideal for what is practical.

See also *advertising effectiveness; advertising media selection; brand equity; perception of brand equity; brand growth strategy; brand strategy; brand value; communications budgeting; market-based assets; marketing metrics; marketing mix; marketing planning; marketing strategy; marketing strategy implementation; marketing strategy models; the advertising budget*

Bibliography

Belch, G.E. and Belch M.A. (2009) *Advertising and Promotion: An Integrated Marketing Communications Perspective*, 8th edn, McGraw-Hill, New York.

Farris, P. and West, D.C. (2007) A fresh view of the advertising budget process, in *The SAGE handbook of advertising*, (eds G.J., Tellis and T., Ambler), SAGE, London, pp. 316–332.

Kissan, J. and Richardson, V.J. (2002) Free cash flow, agency costs and the affordability method of advertising budgeting. *Journal of Marketing*, **66**, 94–108.

Prendergast, G., West, D., and Shi, Y. (2006) Advertising budgeting methods and processes in China. *Journal of Advertising*, **35**, 165–176.

Synodinos, N.E., Keown, C.F., and Jacobs, L.W. (1989) Transnational advertising practices: a survey of leading brand advertisers in fifteen countries. *Journal of Advertising Research*, **29**, 43–50.

consumer rebates: current issues and research

Tim Silk

Overview

Consumer rebates have become a popular price discrimination and promotional tool among retailers and manufacturers to increase sales on a wide variety of durable and consumer goods. The promotion industry estimates that over 400 million rebate checks are distributed annually in the United States totaling $6 billion (Spethman, 2006). Rebates are typically offered in one of two formats. Instant rebates are less common and mirror traditional price discounts by offering a discount at the point of sale. Delayed payment rebates (e.g., mail-in rebates) are more common and require consumers to exert effort after purchase to receive a discount. The effort required can vary but typically involves mailing an application form, sales receipt, and proof of purchase by an expiry date that can range from 15 to 90 days after purchase. On-line rebate redemption is becoming more popular and similarly requires consumers to verify purchase information to protect firms against fraudulent applications. Rebates are distinct from coupons and other forms of price promotion because the effort required to receive the discount occurs after rather than before purchase. This difference has important implications for consumer behavior and rebate redemption rates.

Benefits to Firms and Consumers

Rebates can be effective at increasing sales by encouraging trial, brand switching, and repeat purchase (Promotion Marketing Association, 2005; Tat, Cunningham, and Babakus, 1988). Rebates are attractive to firms because the price discount is paid after purchase, which offers two key benefits. First, rebates have a more favorable impact on key performance metrics such as revenues and comparable store sales vis-à-vis coupons and discounts which are subtracted from the purchase price at the time of sale. Second, firms can benefit from "slippage" which occurs when consumers are attracted by the rebate to make a purchase but later fail to redeem the rebate. Rebates are attractive to consumers because they offer discounts that are typically larger than those of other types of price promotions. Low redemption rates conceivably allow firms to offer larger discounts than would be possible if the discount were realized by all consumers (Chen, Moorthy, and Zhang, 2005). Thus, rebates are essentially a mechanism used to distribute discounts to consumers who perform the effort required to achieve the discount.

Controversies and Research Findings

It has been well documented that consumers often fail to redeem rebates, resulting in low redemption rates (Jolson, Weiner, and Rosecky, 1987; Silk, 2004; Soman, 1998). Slippage and the question of why some consumers fail to redeem rebates has become a controversial issue that has attracted the attention of law makers and researchers. While no federal regulation exists for rebates, state legislators have become proactive in introducing nearly 50 bills on rebates Silk and Pechmann, 2009). Some bills have been enacted into law, resulting in a patchwork of state regulations that complicate national promotions. Research has examined rebate characteristics and consumer behavior that contributes to slippage. This research has shown that foregoing a rebate can be a rational choice that does not disadvantage consumers. For example, Chen, Moorthy, and Zhang, (2005) show that rebates are similar to option contracts where consumers decide after purchase whether or not to exercise their option to redeem, in effect customizing price according to each consumer's marginal cost of time and utility of income. Consumers may also forgo the rebate if it has no bearing on their purchase decision, which is supported by data showing that redemption rates are significantly higher when calculated as a percentage of incremental sales (i.e., consumers influenced by the rebate) rather than total sales (Promotional Marketing Association, 2005). Research also suggests that biases in judgment and decision making also contribute to slippage. Soman (1998) finds that consumers tend to underestimate the cost of future effort relative to apparent savings. As a result, purchase decisions are more sensitive to the rebate amount than to the

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

required effort, whereas redemption behavior is more sensitive to effort than to the rebate amount. Consumers are also highly influenced by redemption deadlines (Silk, 2008). Longer deadlines make consumers more confident that they will redeem, which encourages purchase. Paradoxically, longer deadlines also foster procrastination, which results in lower redemption rates. Consumers are most likely to redeem when deadlines are shortest because short deadlines mitigate procrastination and forgetting. Many manufacturers and retailers have moved toward consumer-friendly rebates in an effort to simplify the redemption process and encourage repeat patronage.

Bibliography

Chen, Y., Moorthy, S., and John Zhang, Z. (2005) Research note – price discrimination after the purchase: rebates as state-dependent discounts. *Management Science*, **51** (7), 1131–1140.

Jolson, M.A., Weiner, J.L., and Rosecky, R.B. (1987) Correlates of rebate proneness. *Journal of Advertising Research*, **27** (1), 33–43.

Promotional Marketing Association (2005) *Mail-In Rebate Benchmarking Study*, Promotion Marketing Association Educational Foundation, New York.

Silk, T. (2004) Examining purchase and non-redemption of mail-in rebates: the impact of offer variables on consumers. Subjective and objective probability of redeeming. Doctoral Dissertation. University of Florida.

Silk, T. (2008) *Getting Started is Half the Battle: The Influence of Deadlines and Consumer Effort on Rebate Redemption*, University of British Columbia, Vancouver, BC, Canada. Working Paper.

Silk, T. and Pechmann, C. (2009) *A Multi-Party Analysis of Consumer Rebates: Controversies and Research-Based Policy Recommendations'*, University of British Columbia, Vancouver, BC, Canada. Working Paper.

Soman, D. (1998) The illusion of delayed incentives: evaluating future effort-money transactions. *Journal of Marketing Research*, **35** (4), 427–437.

Spethman, B. (2006) The Real Problem with Rebates. Promo Magazine (January 11).

Tat, P.K., Cunningham, W.A. III , and Babakus, E. (1988) Consumer perceptions of rebates. *Journal of Advertising Research*, **28** (4), 45–50.

advertising message appeals

Marla B. Royne

An advertising message appeal is the connection between the emotion or logic of the advertisement and the consumer's response to the ad (Wells, Moriarty, and Burnett, 2006). For an advertisement to be effective, the most appropriate appeal must be selected and executed in a creative manner. In general, message appeals can be categorized into two overarching categories: emotional and rational.

Emotional appeals are directed primarily at an individual's psychological, social, or symbolic needs (Arens, Weigold, and Arens, 2008). They are grounded in one's feelings and are generally designed to increase the overall liking for a product, although both negative and positive emotional appeals can be used. The basic objective of an emotional appeal is for consumers to make decisions based on their feelings, as opposed to utilizing logical, rational decision-making approaches. Emotional appeals touch on a variety of feelings including fear, humor, love, anxiety, and sex. To potentially enhance their persuasive effects, an advertisement can actually include more than one type of emotional appeal. Since they are designed to persuade by manipulating one's feelings, emotional appeals often come under scrutiny.

Rational appeals are generally directed at an individual's practical and functional need for a particular product (Arens, Weigold, and Arens, 2008). They are grounded in traditional models of information processing where consumers are expected to make purchase decisions based on an individual's sense of logic and reasoning. The purpose of a rational appeal is to emphasize the characteristics, features, and benefits of the product to demonstrate the value of owning or using that particular brand. Such appeals may include detailed information about a product such as specific features, durability, and value. Successful rational appeals rely on their persuasive power of arguments or reasoning. Examples of rational appeals include price, quality, performance, reason why, and comparative advertising.

For years, effectiveness of the appeal has been believed to be based on the corresponding category of product (Stafford and Day, 1995), whereby there is an appeal by product interaction (e.g., Golden and Johnson, 1983; Aaker, Batra, and Myers, 1992), and that the type of appeal should match that of the product (Holbrook and O'Shaughnessy, 1984; Johar and Sirgy, 1991). More recent research suggests that the choice of message appeal is not that simple. In fact, one perspective is that the most successful advertising integrates functional benefits with emotional values.

Regardless of the type of appeal selected, advertisers must carefully choose the most appropriate one for the particular product being advertised. This means that the characteristics and nature of the product must be considered to ensure that the tone of the advertisement accurately reflects the intended purpose and usage of the featured product.

See also *comparative advertising; copy test methods to pretest advertisements; integrated marketing communication; integrated marketing communication strategy; positioning analysis and strategies*

Bibliography

Aaker, D.A., Batra, R., and Myers, J.G. (1992) *Advertising Management*, 4th edn, Prentice-Hall, Inc., Englewood Cliffs.

Arens, W.F., Weigold, M.F., and Arens, C. (2008) *Contemporary Advertising*, 11th edn, McGraw-Hill Irwin, New York.

Golden, L.L. and Johnson, K.A. (1983) The impact of sensory preferences and thinking vs. feeling appeals on advertising effectiveness. *Advances in Consumer Research*, **10**, 203–208.

Holbrook, M.B. and O'Shaughnessy, J. (1984) The role of emotion in advertising. *Psychology and Marketing*, **1**, 45–64.

Johar, J.S. and Sirgy, M.J. (1991) Value-expressive versus utilitarian advertising appeals: when and why to use which appeal. *Journal of Advertising*, **20**, 23–33.

Stafford, M.R. and Day, E. (1995) Retail services advertising: the effects of appeal, medium, and service. *Journal of Advertising*, **24**, 57–71.

Wells, W., Moriarty, S., and Burnett, J. (2006) *Advertising Principles and Practice*, 7th edn, Pearson Prentice Hall, Upper Saddle River.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

regulating advertising in the United States

Jef I. Richards

The Implications of Communication-as-product

Like any industry, the industry that creates and delivers advertising is subject to a long list of regulations, including those involving employment, zoning, securities, taxation, and more. But the regulations unique to this industry deal largely with its products: the advertisements and related promotional devices. Other industries, of course, also must deal with regulations on their products. The food industry, for example, must meet requirements concerning the quality of ingredients in a product, the amount of time it can remain on a store shelf, and so on. But the advertising industry differs from most others, in that its products are messages, that is, communication.

The First Amendment of the United States (US) Constitution guarantees freedom of communication (i.e., "speech"). It does not, however, promise freedom without some concomitant responsibility. This means there are some limitations imposed on communication, in spite of the amendment's facial assurance of "no law . . . abridging the freedom of speech." For instance, there are "reasonable time, place, and manner" restrictions applicable to virtually any form of communication, including advertising (Langworthy, 1983).

In addition, historically, some different types of speech receive different degrees of protection. Political speech is seen as the principle or "core" purpose behind the first amendment, so it has received more protection than other forms of speech (Richards, 1996). Advertising falls into a category called "*commercial speech*," that has traditionally received far less protection than political speech.

Although the first amendment became law in 1791, it was more than 150 years before an advertising case reached the Supreme Court. And when it did, in *Valentine vs. Chrestensen* (1942), the Court clearly stated that commercial advertising was not constitutionally protected. Indeed, commercial speech was not considered a form of speech protected by the first amendment until *Virginia State Board of Pharmacy vs. Virginia Citizens Consumer Council* (1976). Even after that, the precise extent of that protection continued to be subject to significant debate, even within the Supreme Court (Richards, 1997).

The latest, and in some respects the clearest, statement regarding the status of commercial speech is found in *44 Liquormart vs. Rhode Island* (1996). There, the Court stated that more regulation of commercial speech, as opposed to other speech, is permitted *only* to assure fair bargaining in the marketplace. It suggested that restrictions on commercial speech unrelated to fair bargaining would be subject to standards more like those applied to other protected forms of speech. That decision seemingly elevated the constitutional protections afforded to advertising, and all forms of commercial speech.

As a consequence, not every desired or attempted regulation of advertising will pass constitutional muster. Even prior to *44 Liquormart*, attempts to ban or severely restrict certain types of advertising, like that for cigarettes or alcohol, frequently were thwarted by the first amendment. For those laws, the government must prove that a legitimate purpose will be served in some significant way, and that the law is narrowly tailored to serving that purpose (Central Hudson, 1980).

On the other hand, regulations dealing directly with fair bargaining issues, such as advertising deception, generally found no obstacle in the Constitution (*44 Liquormart*, 1996). Advertising that misleads threatens not only the individual consumer but also the assumptions on which the free market system is founded, so the courts recognize a need to curtail deceptiveness.

Deceptive Advertising Basics

The principle US government agency charged with regulating deceptive advertising is the Federal Trade Commission (FTC). The basis of its authority is the FTC Act (1914), which states, "Unfair methods of competition in or affecting commerce, and unfair or deceptive acts or practices in or affecting commerce, are hereby declared unlawful." By that simple statement, the FTC is granted power over both deceptive

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

and "unfair" acts or practices, which actually encompasses more than deception and reaches marketing practices beyond just advertising. Deceptive advertising, however, is a significant part of the FTC's responsibility (Richards, 1990).

The fundamental principle of deception is that the consumer understands the advertisement to claim something about a product or service that is untrue. The consumer's understanding is key. If the meaning conveyed to the consumer's mind by an advertisement differs from the reality of the product attribute, the advertisement is considered deceptive (Preston and Richards, 1986).

Advertisers might argue they had no intent to deceive, but since consumers can be deceived even under the best of intentions, the FTC considers advertiser intent essentially irrelevant. And because it is the conveyed meaning that is critical, the actual wording in the advertisement can even be true yet still deceive. This means disclaimers in small print will not exempt an advertiser from FTC action, if the consumer is unlikely to read or understand the disclaimers. Again, the consumer's understanding is the governing criterion.

The FTC considers an advertisement deceptive if it meets three requirements:

1. There is a representation, omission, or practice.
2. It is likely to mislead consumers acting reasonably under the circumstances.
3. It is material.

The first of these makes it clear that not only a representation but also an omission of information can result in legal action. Failure to disclose something can be as deceptive as an outright lie. The second, by including the word "likely," reflects the FTC standard that a claim can be regulated without it having deceived anyone, so long as deception is probable. Also, in this second part, the FTC makes it clear that only consumers who act reasonably will receive protection, and that the agency will not go out of its way to protect those holding unreasonable beliefs (Cliffdale Associates, 1984).

The third requirement concerns the claim's "materiality," or importance. A claim (or omission) must not only have the probability of deceiving consumers but it must also be nontrivial. A claim is material to consumers if it is "likely to affect their choice of, or conduct regarding" a product or service (Cliffdale Associates, 1984).

This definition sets the parameters of proof for any advertisement to be declared deceptive.

Evidence. If the FTC receives a significant number of complaints about any advertisement, the agency's staff then collects evidence to evaluate whether it might be deceptive. If the evidence supports a charge against the advertiser, the staff takes that evidence to the commissioners and recommends that a formal complaint be sent to the advertiser. A complaint is the initiation of legal action.

The FTC tends to assume that any claim made by an advertiser is material. This means the only evidence needed to find deceptiveness is to discover (i) what claim is conveyed to consumers' minds and (ii) whether that conveyed claim is accurate.

Commissioners look to a variety of evidence types in determining what claim is being conveyed (Thompson Medical Co., 1984). The most common is to look at the advertisement itself. Where a claim is explicit, no further evidence of the conveyed message may be needed. On the other hand, where it is an implied claim the commission may look to a broad range of evidence, including expert testimony and consumer surveys.

Proof of a claim's accuracy is called "*substantiation.*" For advertisers, the most important aspect of the FTC approach to substantiation is that an advertising claim must not be made without – at a minimum – a *reasonable basis* for believing the claim is accurate. There was a time when advertisers could make a claim and not worry about its accuracy unless and until the FTC challenged the veracity of the claim. But, today the advertiser must have in hand some reasonable evidence of the claim's truth at the time the claim is included in an advertisement. What evidence constitutes a reasonable basis for believing a claim to be true will vary greatly from situation to situation, depending on the product, the type of claim, what experts believe

is reasonable, as well as other factors (Thompson Medical Co., 1984).

The legal burden of proof is on the advertiser. In practical terms, this means when advertisers are charged with making deceptive claims, they are guilty until proven innocent.

The FTC, besides judging whether an advertising claim is deceptive, also plays other roles in advertising regulation.

Other Responsibilities of the Federal Trade Commission

Deceptiveness is not the only way advertising can potentially harm the public. Neither is it the only concern of the FTC. The law empowering the agency gave it authority over both deceptive and "unfair" practices (FTC Act, 1914). This allows the commission occasionally to reach advertising and marketing methods that do not deceive yet might be otherwise harmful. Unfairness is carefully defined in the FTC Act, but it remains sufficiently broad so as to reach a variety of marketing practices by applying a cost-benefit analysis to assess whether the public risks outweigh the benefits of the particular practice (Calkins, 2000).

Historically, the unfairness power has been used sparingly, but it has proved especially useful in reaching practices that threaten particularly vulnerable populations like children and the elderly, where a marketing tactic might not be deceptive and yet takes advantage of a group's susceptibility. More recently, it has been applied to problems arising from new technologies, such as "pagejacking" on the Internet (Calkins, 2000).

This agency also has responsibility for issuing guidelines and regulations. The former are not law, but provide marketers with guidance on how to comply with the law. By contrast, regulations are law. They identify specific practices that already have been determined deceptive or unfair, or to violate some other law enforced by the FTC, and noncompliance with a regulation can result in a fine or other penalty.

FTC Remedies

Penalties can result for violation of regulations, as well as for violations of FTC orders. However, the term "*penalty*" connotes punishment, which is beyond the FTC's power. Instead, its authority is limited to "remedies," orders designed to force a marketer to repair a situation rather than to punish the marketer.

For example, once a marketer is found responsible for using a deceptive claim, the commission has only a few options, none of which constitute punishment. A "cease and desist order" is the most common remedy. This simply orders the marketer to stop using that claim. If the deceptiveness stems from important information being absent from the advertisement, another possible remedy is "affirmative disclosure," which commands the advertiser to include certain information in future advertisements. And the most rare remedy is "corrective advertising," which is reserved for only those situations where there is a long history of using a deceptive claim. In those situations, the FTC can require a more explicit corrective disclosure (Wilkie, McNeill, and Mazis, 1984).

It should be noted, however, that the FTC most often does not even reach the point of ordering any of those remedies. Most cases end long before the advertisement is officially found deceptive, because advertisers merely acquiesce to FTC demands to terminate or modify the offending claim before any hearing is held. This results in a "consent decree," whereby, the advertiser voluntarily consents to make changes.

And then there are other particularly egregious cases where the FTC's available remedies are just inadequate to fix the problem. While it has no authority to punish marketers, in these situations, the commission can turn the case over to a court, and courts have far more options available to them, including the ability to order the marketer to pay back consumers for their losses (Ward, 1992). These "section 13(b)" cases have become a popular solution for the regulators.

Although the FTC is the primary federal force to stem the excesses of advertising, it is not, by any means, the only form of advertising regulation. Indeed, the FTC Act is not the only law that imposes limits on the advertising industry.

Other Sources of Regulation

There are scores of laws that affect advertising. Many of them are industry-specific. As just two examples, there are laws restricting

certain claims used in advertising agricultural plant materials, and there also are laws regarding advertising of textile fiber products. There also are laws that cut across multiple industries, like the laws aimed at protecting consumer privacy, which clearly affect advertising.

At the same time, there are federal administrative agencies other than the FTC engaged in advertisement regulation. The two most actively involved are the Federal Communications Commission (FCC), because it regulates media over which so many advertisements are conveyed, and the Food and Drug Administration (FDA), which oversees so many of the products being advertised. Other agencies, like the Securities Exchange Commission or the Department of Transportation, likewise may play a role on the more rare occasions where their area of responsibility is involved.

And then there are state and local governments that frequently do play a role in advertising regulation (Richards, 1991). States have laws, including what are sometimes called "*Little FTC Acts*," to regulate advertising within their borders.

Conclusion

The legal environment of advertising in the United States is complex, because the range of products, services, advertising concepts, and claims in advertising is virtually limitless. Moreover, as with any body of law, it is constantly changing. Consequently, any specific advertisement or claim might invoke legal principles not even mentioned above.

Bibliography

44 Liquormart vs. Rhode Island, 517 U.S. 484 (1996).

Calkins, S. (2000) FTC unfairness: an essay. *Wayne Law Review*, **46**, 1935–1991.

Central Hudson Gas & Electric vs. Public Service Commission of New York, 447 U.S. 557, (1980).

Cliffdale Associates, 103 F.T.C. 110, (1984).

Federal Trade Commission Act, 15 U.S.C. §§ 41-58, (1914), as amended.

Langworthy, E.A. (1983) Note: time, place, or manner restrictions on commercial speech. *George Washington Law Review*, **52**, 127–145.

Preston, I.L. and Richards, J.I. (1986) Consumer miscomprehension as a challenge to FTC prosecutions of deceptive advertising. *The John Marshall Law Review*, **19**, 605–635.

Richards, J.I. (1990) *Deceptive Advertising: Behavioral Study of a Legal Concept*, Lawrence Erlbaum Associates, Hillsdale.

Richards, J.I. (1991) FTC or NAAG: who will win the territorial battle? *Journal of Public Policy and Marketing*, **10**, 118–132.

Richards, J.I. (1996) Politicizing cigarette advertising. *Catholic University Law Review*, **45**, 1147–1212.

Richards, J.I. (1997) Is 44 liquormart a turning point? *Journal of Public Policy and Marketing*, **16**, 156–162.

Thompson Medical Co., 104 F.T.C. 648, 790, (1984), *aff'd*. 791 F.2d 189, (DC Cir. 1986), cert, denied 479 US 1086, (1987)

Valentine vs. Chrestensen, 316 U.S. 52, (1942).

Virginia State Board of Pharmacy vs. Virginia Citizens Consumer Council, 425 U.S. 748, (1976).

Ward, P.C. (1992) Restitution for consumers under the federal trade commission act: good intentions or congressional intentions? *American University Law Review*, **41**, 1139–1197.

Wilkie, W.L., McNeill, D.L., and Mazis, M.B. (1984) Marketing's "scarlet letter": the theory and practice of corrective advertising. *Journal of Marketing*, **48**, 11–31.

direct-to-consumer (DTC) advertising

Marla B. Royne

Direct-to-consumer (DTC) advertising is the advertising of pharmaceutical drugs and other prescription products directly to the consumer. While the first modern DTC ad appeared in 1981 (Friedman and Gould, 2007), a voluntary moratorium on such advertising followed (Woloshin *et al*., 2001). In 1985, the moratorium was lifted, but the information requirements were so strict that advertisers often found DTC advertising to be cost prohibitive. In 1997, the Food and Drug Administration (FDA) relaxed the existing regulations, paving the way for massive national advertising campaigns for prescription drugs (Auton, 2004); today, DTC advertising permeates the mass media. In 2008, the pharmaceutical industry spent $4.3 billion on advertising spending (en-us.nielsen.com, 2009) Despite this massive budget, it actually represents a decrease of 18.4% from the previous year (en-us.nielsen.com, 2009), Moreover, Americans see as many as 16 hours of prescription drug advertising each year (Brownfield *et al*., 2004).

DTC advertising has been a particularly controversial topic since the modified legislation in 1997 that generated large-scale television and print advertising campaigns for prescription drugs (Auton, 2004). And while the Pharmaceutical Research and Manufacturers of America (PhRMA) subsequently issued voluntary guidelines on DTC advertising, questions abound on whether regulations adequately protect consumers from such overt marketing communications.

The United States and New Zealand are currently the only countries that actually allow DTC advertising, although countries in Europe have seen increased pressure to permit the practice. At the same time, New Zealand's health officials have been lobbying for a ban on DTC advertising for several years (Advertising Age, 2005). The controversy stems from the debate on whether DTC advertising *informs* or *influences* customers. Individuals supporting DTC advertising argue that media messages provide the information needed for consumers to assist their doctors in making an informed choice in their own care (Auton, 2004; Meek, 2001). Those individuals against the practice feel that DTC advertising is not informative and exists solely to persuade customers to ask their doctors for an advertised prescription. Critics further argue that marketing tactics serve only to increase the demand for more expensive medications (Peyrot *et al*., 1998).

In addition, DTC advertising critics believe DTC advertising medicalizes normal everyday experiences and treats these experiences as a disorder (Mintzes, 2002). Another concern is that ads sometimes leave the impression that the newer advertised drugs are superior to existing older drugs, when in fact this is often not the case. Critics also suggest that advertisers fail to include "balanced" information in their ads, one of the guidelines recommended by the 2005 PhRMA voluntary guidelines (Welcher, 2005). Interestingly, an increasing number of consumers do not feel that such advertising helps them make better health decisions nor do they feel that it provides enough information on the risks or potential negative effects of the drug being advertised (Friedman and Gould, 2007).

While several DTC advertisements have been pulled due to misleading information, others have been lauded for providing the information needed in a positive, upbeat manner (e.g., Dick-Rath, 2009). This campaign for Gardisal, a vaccine developed to prevent human papillomavirus (HPV), had positive outcomes for both the consumer and the company (Dick-Rath, 2009). This type of campaign demonstrates that DTC advertising may, indeed, have an important place in the consumers' decision-making process when it comes to their health.

Despite the vast concerns about DTC advertising, it is unlikely that the United States will ever see a complete ban on the practice. However, with PhRMA's voluntary guidelines and the FDA's oversight, changes to DTC advertising may be on the horizon. Such changes should result in positive outcomes for consumers, first, and pharmaceutical companies, second.

See also *regulating advertising in the United States*

Bibliography

Advertising Age (2005) New Zealand to ban DTC advertising by '06. *Advertising Age*, **76** (36), 1–2.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Auton, F. (2004) The advertising of pharmaceuticals direct to consumers: a critical review of the literature and debate. *International Journal of Advertising*, **23**, 5–52.

Brownfield, E.D., Bernhardt, J.M., Phan, J.L. *et al.*, (2004) Direct-to-consumer drug advertisements on network television: an exploration of quantity, frequency and placement. *Journal of Health Communication*, **9** (6), 491–497.

Dick-Rath, D. (2009) DDR on DTC. *Medical Marketing and Media*, **44** (4), 26.

en-us.nielsen.com (2009) http: //en-us.nielsen. com/main/news/news_releases/2009/march/u_s__ad_spending_fell.

Friedman, M. and Gould, J. (2007) Consumer attitudes and behaviors associated with direct-to-consumer prescription drug marketing. *Journal of Consumer Marketing*, **24** (2), 100–109.

Meek, C. (2001) *Direct to Consumer Advertising of Prescription Medicines: A Review of International Policy and Evidence*, Royal Pharmaceutical Society of Great Britain.

Mintzes, B. (2002) For and against: direct to consumer advertising is medicalising normal human experience. *British Medical Journal*, **324** (7342), 908–909.

Peyrot, M., Alperstein, N.M., Van Dore, D., and Poli, L.G. (1998) Direct-to-consumer ads can influence behavior. *Marketing Health Services*, **18** (2), 26–32.

Welcher, J. (2005) Drug advertising guidelines draw criticism. *Managed Healthcare Executive*, **15** (9), 5.

Woloshin, S., Schwartz, L.M., Tremmel, J., and Welch, H.G. (2001) Direct to consumer advertisements for prescription drugs: what are americans being sold?. *The Lancet*, **358** (6), 1141.

product placement

Cristel Antonia Russell

Product placement refers to promotional practices that integrate brand communications within the content of entertainment products. Brand communications are now present in the content of a broad range of entertainment vehicles, including TV and movies (Wasko, Phillips, and Purdie, 1993), radio shows, songs and music videos, video games, plays, and even novels (Friedman, 1985). The increased mingling of advertising with the entertainment world has generated a slue of newly coined terms to reflect these trends, such as hybrid advertisement (Balasubramanian, 1994) or the "Madison and Vine" expression, reflecting the physical intersection of the advertising industry's New York City hub, on Madison Avenue, and the entertainment hub on Vine Street (Donaton, 2004).

Many factors contribute to advertisers' increased interest in and use of product placement. It has grown mainly in reaction to the increasing advertising clutter, escalating advertising costs, and the reduced effectiveness of traditional advertising messages. Consumers are exponentially exposed to commercial messages but at the same time they are finding new ways to avoid them. An In-Stat/MDR survey found that 54.3% of consumers claim to skip 75–100% of commercials (Jaffe, 2005). In 2004, a Knowledge Networks study concluded that 47% of viewers switch channels while watching TV (Jaffe, 2005). These statistics do not account for the impact of new technological advances which are giving consumers more control over how they consume entertainment, and making it increasingly easy for them to avoid commercial messages. With the introduction of personal video recorders (PVRs), also referred to as digital video recorders (DVRs), such as TiVo or replay TV, consumers can not only more easily fast-forward through commercials but they can now also easily skip them altogether with a PVR's auto-skip function. A Forrester Research's study of PVR usage by 588 users in the United States found that 60% of their time, on average, was spent watching programs that were prerecorded or delayed, which in turn resulted in 92% of commercials being skipped. Thirty percent of respondents said they watched no commercials at all (Zutter, 2005).

Product placement epitomizes the blurring of the lines between advertising and entertainment. In audiovisual media, placements can be categorized based on their modality of presentation, visual or audio. The visual appearance of brands is justified by the need to create realistic settings for movie or TV sets. According to records available on www.brandchannel.com, an average of 18 brands was visually placed in the roughly two hundred top movies released between 2001 and 2005. During the 2004–2005 TV season, more than 100 000 product placements appeared on the six broadcast networks, an increase of 28% from the previous season, according to Nielsen Media Research (Manly, 2006). According to Nielsen Product Placement research, the most common form of product placement in 2005 was a simple spot in the background of a show or the brand used as visual prop in the program (Edwards, 2006a). Visual placements are common in video games, where they usually take the form of hoardings, billboards, and traditional signage placed throughout a game, for example, on the sides of buildings in a motor sport game, which essentially mimic real-life advertising spaces (Murphy, 2006). Audio mentions tend to have more impact because they imply an endorsement by the celebrity. For instance, the popularity and reputation of many talk show hosts can make a simple mention of a product on a talk show extremely successful.

A Historical Perspective

Although product placement is gaining in popularity, it is not new, and there is a long history of such intermingling between content and advertising. Marketers' approach to using entertainment content to promote their products dates back to the use of branded products in motion pictures (Newell, Salmon, and Chang, 2006). This practice was variously termed *publicity by motion picture* (Dench, 1916), *co-operative advertising* (Harrower, 1932), *tie-in advertising* Tie-in Advertising (1951) and *trade outs* (Lees and Berkowitz, 1978), or even *exploitation* (Steele, 1925). It represented a

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

cooperative venture between a media maker and a manufacturer, in which on-screen exposure of a product, off-screen endorsement by an actor, or a combination of those were traded for paid advertising and unpaid promotions by the manufacturer (Eckert, 1978). The use of tie-ins became standardized in the 1930s. The Walter E. Kline Agency in Beverly Hills provided studio executives with multiple-page lists of products available for on-screen use in motion pictures, including Remington typewriters, IBM tabulating machines, Singer sewing machines, and General Electric appliances (Newell, Salmon, and Chang, 2006). Products were offered rent free in return for publicity stills for use in manufacturers' advertising (Kline, 1931).

These tie-in promotions benefited multiple players in the filmmaking and distribution industry. For the motion picture producer, they delivered free props. For the manufacturer, they provided in-theater exposure for products to a captive audience and the chance for the product to be associated with well-known actors on- and off-screen. In addition, they provided the marketing option to use footage or publicity stills from the movie in the company's outside advertising. For the motion picture distributor, this "as seen in the motion picture..." early form of cross-promotion provided zero-cost advertising for the motion picture. Many of the same benefits apply today.

The practice eventually evolved into sponsor-owned programs, where manufacturers had full control over how their products were portrayed (*see* SPONSORSHIP AND EVENT MARKETING). Starting in the 1930s, consumer product manufacturers invested in the production of radio programs to reach their target audiences (Lavin, 1995). This phenomenon was particularly visible in the "soap opera" genre, a term that actually testifies to the blending of advertising of actual soap products and programming. Similarly, radio programs were developed directly by detergent companies, notably Procter & Gamble, to promote their brands by integrating them into the scripts. When soaps emerged on TV in the 1950s, the close connection between programming and advertising continued and marketers maintained direct control over the shows' storylines and creative design (Barnouw, 1975). Advertising agencies produced shows like "The Colgate Comedy Hour" and "Texaco Star Theater," in which a chorus line of dapper gas-station attendants opened each show by singing the Texaco jingle ("Oh, we're the men of Texaco, we work from Maine to Mexico") before introducing the host, Milton Berle (Manly, 2005).

The practice of show sponsorship began to decline as advertisers realized they could better reach their target markets by spreading their advertising budgets across many shows rather than by spending it all on one (Savan, 1996). Since the mid-1980s, however, the trend has reversed and advertisers are now again trying to tie their brands to entertainment vehicles. One factor that fostered the unprecedented proliferation of product placements and campaigns tied to movies occurred during the 1980s and early 1990s (Wasko, Phillips, and Purdie, 1993) was the use of Reese's Pieces to attract the alien in the 1982 hit movie *E.T.* The success of the movie and the placement led to an increase in awareness of the brand as well as sales increases of 65% (Winsky, 1982).

How the Product Placement World Works

A direct consequence of the increased popularity of product placement is the boom of a specialized industry (Wenner, 2004). Product placement is practiced by more than 100 companies internationally (Nelson, 2004). The institutionalization of the industry is evidenced by such professional trade associations as the Entertainment Marketing Association (EMA) whose association members include placement agents, studio representatives, and marketers. A review of EMA membership reveals that in North America, product placement firms are mainly located in the Los Angeles area, hub of the entertainment world, and in major advertising centers such as New York, Chicago, and Toronto Entertainment Marketing Association (2007). The three main types of players are the production side, the agency side, and the client side.

The production side. These are the producers of entertainment content, such as TV and movie

studios, music production companies, or video game makers. In the audiovisual arena, the bases for identification of potential for inclusion of product placements are scripts. Studios send scripts to clients whose permissions are legally required to use these brands in their scenes, either directly or through the intermediary of product placement agencies. Movies generally include between 100–150 brands that need such permissions (Russell and Belch, 2005). Decisions on whether to include a branded product and choices of which brand to include are sometimes made by the directors, producers, or actors. Their criteria for selection are generally driven by artistic needs or the desire to make the story realistic.

Increasingly, an objective for using product placements is financial. This may come in various forms. Federal Communications Commission (FCC) rules prohibit payment for placements on TV shows but not in other media, such as movies. However, even on TV, the regulations do not prevent other forms of agreements between producers and advertisers. A common form involves the provision of the product as payment, similar to trade deals made in the media industry, that is, providing products and services in return for having the product shown at no direct cost. This form of payment reduces production and overhead costs.

A major limitation of simply sending products out to a placement agent or directly to the producers is that it provides limited to no control over how the brand is ultimately portrayed. Of course, producers must take into account existing contracts with advertisers to avoid conflict of interests but this is always clear-cut. Another concern on the production side is the potential violation of creative and artistic freedom associated with increased commercial content (Manly, 2005). Producers sometimes turn down placement deals (Manly, 2005). Some creators of television shows worry that the commercial impulse could turn storytellers into shills and keep more provocative fare off the big networks.

The client side. The companies that use product placements to promote their products and/or brands range from small to large Fortune 100 corporations. Clients sometimes review scripts provided by their agents or directly by the production companies. Their perspectives often differ from agents' because they have to consider the placement as part of their overall brand strategy. However, given the relative novelty of the practice, it is not always clear who in the client organization has responsibility for it. In fact, organizations that do not think of product placement as a fully integrated component of their strategic marketing activities, but rather as an accessory tool, often delegate such responsibilities primarily to their agents.

The agency side. Analogous to advertising agencies, placement agencies serve as the intermediary between the production side and the client side. These agencies are paid a fee for their services and essentially become agents for the producers (Bohn, 1986). The primary basis for the agency participation in this industry is to receive direct monetary compensation. Many of today's placement agents started as prop masters on the studio side and realized the opportunity for an intermediary to streamline and improve the communication channels between the two sides, and hence the many agency start-ups in the 1980s.

Some companies now coordinate their product placements in-house but many use an intermediary agency that has closer links to the entertainment world. In many instances, agencies are provided with advance copies of these scripts which are reviewed for placement opportunities, matching existing clients to these opportunities, and/or seeking new participants. Because placement agents are not trained as brand managers, their criteria for deciding whether one of their clients' brands is appropriate for a particular medium is often more opportunistic than strategic.

The product placement industry is evolving from a simple means for studios to cut costs, and marketers to expose their products to a more sophisticated industry in which product placement agencies are increasingly assuming the role of traditional advertising agencies. At the same time, traditional advertising and media agencies are beginning to incorporate product placement in their services. Major holding companies are increasingly adding entertainment consultants and public relations powerhouses to their portfolios.

Role of Product Placement in Integrated Marketing Communications (IMC) Mix (*see* integrated marketing communication strategy)

Placement efforts can have a strategic role in a fully integrated promotional campaign. There is no question that developing objectives and evaluating the effectiveness of placements is easier when they are associated with other, more familiar components of the integrated marketing communications (IMC) mix. Whenever a placement is used to initiate a theme for a more traditional advertising campaign, advertising objectives and effectiveness measurement tools can be used. Since cross promotions are becoming more common to enhance product placements and integrations in movies, efforts can be assessed with sales promotions methods. Placements can also serve as the backbone of a public relations campaign, to generate word of mouth or press coverage. For instance, in what was named a "Top 5 Marketing Events in 2004," Oprah gave away 276 Pontiac G6 cars to car-deprived audience members (Grillo, 2005). The stunt cost GM at least $7 million but resulted in more than 600 TV news stories, and the words "Oprah" and "Pontiac G6" rose to the ranks of Yahoo! and Google's top 10 most requested search terms (Grillo, 2005). In the end, the placement-based PR campaigns resulted in Pontiac selling more than 16000 G6 models in 2004 (Grillo, 2005). Finally, the increasing opportunities with interactive television and Internet-based campaigns allow the use of direct marketing tools to evaluate the impact of placements, especially when a direct response component is built into the campaign.

Measuring Impact (*see* marketing metrics; advertising effectiveness)

One of the main difficulties related to product placement lies in measuring impact. Key indicators such as audience reach and frequency are often next to impossible to obtain (Murphy, 2006). Some argue that the minimal cost of stand-alone placements does not necessarily warrant much attention be paid to the returns (Schiering, 2003). As the industry matures and opportunities for large scale placements and integrations grow, valuation systems will be needed. However, so far, there is little sophistication in this realm and value-like effectiveness-assessment is essentially still very much subjective and anecdotal evidence prevails, mainly with a biased focus on success stories (Russell and Belch, 2005). In a 2005 Association of National Advertisers survey of marketers, 40% of the survey's respondents thought there ought to be an industry standard measurement for branded-entertainment deals, but 80% of those involved said they are forging ahead on their own and assessing the impact of their branded-integration efforts themselves (Klaassen, 2006).

Product placement may be one of the few areas where academic research and practitioners' knowledge are at the same level. Despite recent academic research addressing the effects of product placement on consumers (Astous and Chartier, 2000; Russell, 2002), industry practitioners have only recently started to develop tools to systematically assess placements' effectiveness. Many in the industry have been reluctant to adopt any such measures. Product placement agencies do not advocate any form of measurement, thereby avoiding being held accountable for their performances (Russell and Belch, 2005). Many agents simply provide clients with recordings of the final cuts which display the brand as evidence of their work. Some clients have been convinced that it is not possible to get accurate measures; others are simply satisfied with impressions and/or ego involvement objectives.

A number of firms are competing to develop key measurement formulas in an attempt to establish an industry standard. While many of these (Nielsen Media Research, Deutch/iTVX, among others) focus on the monetary value of the placement and/or media equivalencies, others (IAG Research, NextMedium) rely on outcome-oriented assessments including recall, association, and other consumer responses to the placement (Schiller, 2004). Capitalizing on clients' wishes to determine their placements' ROI, companies providing valuation services have proposed a series of factors that should be considered in determining the value of placements. Such factors usually include

placement characteristics (e.g. screen time, character building, etc.), context characteristics (competitive messages, opportunity for distraction), and audience characteristics (size, demographics). Most of the valuation models compute factored impressions as a function of media impressions (Nielsen ratings for TV or box office + video + cable viewership for film) and as a function of the "level" of the placement itself. Such levels are determined based on exposure and potential for recall with more prominent placement and/or auditory placements perceived as better.

Different metrics for measuring effectiveness should be used depending on the objective. The dependent variables proposed range from awareness and recall to demonstrated purchase behaviors. The potential role of placement in the communications strategy depends on the level of integration of the brand in the entertainment content. Whereas simple product placements are often treated as ad hoc bonus branding efforts, more complex branded-entertainment efforts can become the centerpiece of an integrated marketing communications campaign, leading directly to sales. The hierarchy of effects theory provides a useful framework to link placement objectives with their likely impact on consumers and to identify what metric to use to measure that impact. The theory consists of three stages, each focusing on different types of impact on the audience. The cognitive stage focuses on developing awareness and simple product placements, even on their own, can generate exposure for brands. The affective stage focuses on attitudinal effects, the development of positive brand attitudes. Finally, the behavioral stage focuses on actions and purchases. Behavioral effects can be evaluated more readily when placements are combined with cross promotions, using direct response tools.

Given the difficulty in assessing effectiveness, it is also hard to assign a price to a placement. Some attempts have been made at comparing placements to a traditional TV commercial, but that approach has many limitations because the two promotional tools are very different. Others suggest comparing it to publicity and using tools from the public relations arena.

Strategic Issues

Lack of control. Getting a brand in the content of a TV program or movie often means that the end result is out of the control of marketers or the placement agents. Even expensive placement deals do not guarantee what the final product will look like. A famous case is the Reebok placement in the 1996 movie *Jerry Maguire*. In exchange for Reebok's agreement to plan and execute a $1.5 million promotional campaign, TriStar promised specific in-film placements of the Reebok name and products (Reebok, 1997). When Reebok first read the *Jerry Maguire* script in December 1995, executives loved the story and felt it reflected Reebok's own philosophy that the basic passion for winning creates the possibility to win. In the draft script initially read by Reebok, Tidwell makes certain negative comments about Reebok. For example, the script included a scene where Tidwell complains that he wasn't "getting no love from Reebok" and where he "boils down" his Reebok story to Fuck Reebok. All they do is ignore me .... Always have!" Notwithstanding these negative references to Reebok, the original script contained an uplifting commercial rolling with the end credits. Yet, the commercial was cut and never made it to the final film and, to make matters worse, the film's disparaging comments about Reebok stayed. As a result, Reebok sued Tristar Pictures for breaking a product placement deal and, although in the end, Reebok won, the *Jerry Maguire* movie with its disparaging comments about Reebok remains.

The lack of control is especially common for placements in unscripted TV programs. Because of the nature of reality TV, marketers do not have complete control over how their product is represented in the show and this can be detrimental to the brand. The lack of control is directly related to the nature of the entertainment industry, where producers do not want to be dictated what their content should include. As a result, some companies are developing their own content outside of the main entertainment content, mainly through digital extensions. For instance, Mindshare, the agency in charge of deodorant Degree's advertainment, worked in partnership with Fox's program *24* to create a series of episodes that will be only available

on the web on a Degree-sponsored site. They engaged the show's creative staff and created a new character around which they developed and filmed original stories involving the brand. This ensured that the final content was exactly as suited the brand (Fruitkin, 2006).

*Ethical and legal concerns (*see SELF-REGULATION*;* GLOBALLY INTEGRATED MARKETING COMMUNICATIONS*;* REGULATING ADVERTISING IN THE UNITED STATES*).* Product placement is sometimes referred to as *stealth marketing*, because it blurs the line between publicity, the circulation of messages for free in the hope of further dissemination without attribution of source, and advertising, the paid circulation of messages with attribution (Goodman, 2006). It is viewed as problematic because it often conceals sponsorship although the message is actually paid.

Some controversy has arisen over the potential disguise associated with product placement. Gary Ruskin, executive director of Ralph Nader-affiliated Commercial Alert says that advertainment represents the commercial takeover of every nook and cranny of our lives and culture (Lasswell, 2004). The watchdog group filed complaints with both the FTC and FCC, seeking clearer viewer notification of brand embedding. To date, the government has taken no action.

The distinction between editorial content and advertising is especially problematic in news or semi-news programs, where the content is taken as information and the paid relationship is not disclosed to consumers. The potentially hidden nature of advertainment techniques has raised special concerns for young audiences. Although the FCC has strict rules on advertising in children's TV programs, including a limit on the number of spots per hour and a ban on using show characters in ad breaks, it does not have any rules regarding product placement. Of particular concern is the fact that TV shows featuring products linked to obesity, such as sodas and fast food, are among the most popular with children (Edwards, 2006b). Placements for alcohol brands congregate in shows with audiences that skew heavily toward underage viewers.

Product placement falls in a gray area of regulation. Mainly, this is because today's sponsorship disclosure laws focus on yesterday's technology and fail to operate in the electronic media (Goodman, 2006). The only legislation that can be loosely applied to product placement dates back to when Congress first required broadcasters to identify their sponsors in the 1927 Radio Act. The principle that audiences are entitled to know who is trying to persuade them was continued in section 317 of the 1934 Communications Act. Although the FCC upholds this in rules that say there should be on-screen disclosure of paid-for product placements in TV shows, usually in the end credits, these are not strictly enforced. As a result, advertainment efforts go largely unregulated. The FTC also has regulations in place designed to prevent unfair or deceptive brand communications, especially to children. However, this largely covers advertising and the FTC argues that as product placement does not make "objective claims" about a product, it falls outside these rules and no on-screen advertisement disclosure is required.

Consumer advocacy groups such as Commercial Alert are lobbying for stronger regulation, arguing that the few rules that do exist are systematically violated. They also claim that the line between TV programs and infomercials has become blurred and that product placement is deceptive because it flies under the viewer's skeptical radar. Product placements and script write-ins can be so subtle that viewers fail to identify them as a veiled commercial.

Finally, it is no secret that, in product categories such as alcohol and tobacco, where direct advertising for the product is highly regulated or even illegal, product placements can serve to bypass government or self-imposed industry regulations (Pechmann and Shih, 1999).

The Future of Product Placement

Estimating the monetary size and scope of the placement industry is a difficult task because many product placement deals are unpaid or undisclosed. However, there is no doubt that it is growing and will continue to grow. The overall value of the global product placement market, including the barter/exposure value of nonpaid placements, grew 27.9% to $5.99 billion in 2005

(PQ Media, 2006). The growth of product placement continues to significantly outpace that of traditional advertising and marketing and global paid product placement spending is forecasted to continue growing to reach over $14 billion by 2010 (PQ Media, 2006).

Yet, the relatively high cost of some of the deals combined with the lack of standard measurement tools also points to major limitations to product placement compared to more traditional, and better-understood communication tools. In addition, though the initial rationale was to break through commercial clutter, product placements have now become so prevalent that they too are creating clutter. Instead of being outside the content, clutter is now creeping up inside entertainment content (Edwards, 2006b). This irony has been noted in the advertising industry but, so far, little has been done to limit the amount of product placements in programs. Some executives in the advertainment industry believe that the decline in ratings of heavily branded entertainment programs and the cancellation of some of the major ones, like *The Apprentice*, are in direct proportion to the amount of product placements in them. Although this is a recent and yet a certain trend, some companies, including the major players such as Coca-Cola (Edwards, 2006b), Procter & Gamble, or Unilever (Van Hoffman, 2007), are beginning to react to this new form of clutter by rethinking their marketing initiatives. Product placement will continue to evolve as advertisers look for better and more effective ways to communicate with consumers.

The growth of product placement has led to much speculation and uncertainty over the future of the industry. It has become such a popular tool that there is now talk of reverse placements: brands that would make their debut in an entertainment vehicle and then be introduced to the real world, such as the restaurant chain Bubba Gump Shrimp Company that was created after the 1994 Forrest Gump movie (Wasserman, 2007), or the BMW M5 prototype that was introduced inside a video game to allow people to actually experience the look, the sound, and the handling before there was even a working prototype available (Brand Strategy, 2007).

Bibliography

Astous, d'A. and Chartier, F. (2000) A study of factors affecting consumer evaluations and memory of product placement in movies. *Journal of Current Issues and Research in Advertising*, **22** (2), 31–40.

Balasubramanian, S.K. (1994) Beyond advertising and publicity: hybrid messages and public policy issues. *Journal of Advertising*, **23** (4), 29–46.

Barnouw, E. (1975) *Tube of Plenty: The Evolution of American Television*, Oxford University Press, New York.

Bohn, J. (1986) Business marketers "Go Hollywood". *Business Marketing*, **71**, 21.

Brand Strategy (2007) Reverse Product Placement in Gaming. London, May 9, 2007, p. 28.

Dench, E.A. (1916) *Advertising by Motion Pictures*, Standard, Cincinnati.

Donaton, S. (2004) Cautionary tales from Mad and Vine. *Advertising Age*, **75** (26), 12.

Eckert, C. (1978) The Carole Lombard in Macy's Window. *Quarterly Review of Film*, **3** (1), 1–21.

Edwards, J. (2006a) Placements + commercials = too much of a bad thing? *Brandweek*, **47** (43), 14.

Edwards, J. (2006b) Coke forces TV placements clutter debate into the open. *Brandweek*, **47** (3), 12.

Entertainment Marketing Association (2007) www.emainc.org

Friedman, M. (1985) The changing language of a consumer society: brand name usage in popular american novels in the postwar era. *Journal of Consumer Research*, **11**, 927–938.

Fruitkin, A. (2006) Advanced placement. *Adweek*, **48** (18), SR4.

Goodman, E.P. (2006) Stealth marketing and editorial integrity. *Texas Law Review*, **85** (1), 83–152.

Grillo, J.B. (2005) General Motor's wildest dream. *Broadcasting and Cable*, **135** (4), 56.

Harrower, J. (1932). Exploitation. in Film Daily Yearbook of Motion Pictures 1932, Los Angeles, pp. 404–405.

Jaffe, J. (2005) *Life After the 30-Second Spot: Energize Your Brand with a Bold Mix of Alternatives to Traditional Advertising*, John Wiley & Sons, New York.

Klaassen, A. (2006) Marketers fear being fleeced at corner of Madison and Vine. *Advertising Age*, **76** (13), 3,124.

Kline, W.E. (1931) Correspondence from W. E. Kline to Lincoln Quarberg. Margaret. Herrick Library Special Collections, Lincoln Quarberg Collection, Beverly Hills.

Lasswell, M. (2004) Brand me, baby! *Broadcasting and Cable*, **135**, 34.

Lavin, M. (1995) Creating consumers in the 1930's: Irna Phillips and the Radio Soap Opera. *Journal of Consumer Research*, **22**, 75–89.

Lees, D. and Berkowitz, S. (1978) *The Movie Business*, Random House, New York.

Manly, L. (2005) When the Ad turns into the story line. *New York Times*, **155** (53355), 1.

Manly, L. (2006) Is it a show or commercial? *New York Times Upfront*, **138** (8), 14. New York.

Murphy, J. (2006) Gaming's advertising appeal *Media*, **13**: 1.

Nelson, R.A. (2004) A product placement resource guide: recommended publications and websites. *Journal of Promotion Management*, **10** (1/2), 259–267.

Newell, J., Salmon, C.T., and Chang, S. (2006) The hidden history of product placement. *Journal of Broadcasting and Electronic Media*, **50** (4), 575–594.

Pechmann, C. and Shih, P.C.-F. (1999) Smoking scenes in movies and antismoking advertisements before movies: effects on youth. *Journal of Marketing*, **63**, 1–14.

PQ Media (2006) Product Placement Spending in Media 2006, available at www.pqmedia.com/global-product-placement-2006.

Reebok (1997) Jerry Maguire. US CD California CV 96-8982 SVW http://www.courttv.com/archive/legaldocs/business/reebok.html

Russell, C.A. (2002) Investigating the effectiveness of product placements in television shows: the role of modality and plot connection congruence on brand memory and attitude *Journal of Consumer Research*, **29**, 306–318.

Russell, C. and Belch, M. (2005) Managerial investigation into the product placement industry. *Journal of Advertising Research*, **45** (1), 73–92.

Savan, L. (1996) Your show of shills. *Time*, **147**, 70–71.

Schiering, M. (2003) Red carpet branding. *Brandweek*, September 15, 28–30.

Schiller, G. (2004) Giants, Startups of Placement Valuation. HollywoodReporter.com, December 30, pp. 1–5.

Steele, R. (1925) Exploiters magnificent *Outlook*, July 393–396.

Tie-in Advertising (1951) Consumer Reports, pp. 43–44.

Van Hoffman, C. (2007) P&G, unilever tuning out network product placements *Brandweek*, **48** (20), 10.

Wasko, J., Phillips, M., and Purdie, C. (1993) Hollywood meets madison avenue: the commercialization of U.S. Films. *Media, Culture, and Society*, **15**, 271–293.

Wasserman, T. (2007) Forward thinkers push reverse product placement. *Brandweek*, **48** (5), 5.

Wenner, L.A. (2004) On the ethics of product placement in media entertainment. *Journal of Promotion Management*, **10** (1/2), 101–132.

Winsky, J.M. (1982) Hershey befriends extra-terrestrial. *Advertising Age*, **1**, 66.

Zutter, S. (2005) Big trends. *Marketing Magazine*, **110** (12), 11.

marketing communication on the Internet

Youngseon Kim and Tina M. Lowrey

Viral marketing refers to a marketing strategy or a marketing phenomenon that facilitates and encourages people to pass along a marketing message to other people, generating the potential for exponential growth in the exposure and influence of the message (Wilson, 2005). This relatively new marketing communication strategy seems to have got its label "viral" from its self-duplicating processes, analogous to the spread of pathological and computer viruses. Viral marketing has also been described as "word-of-mouth," off the Internet. However, viral marketing on the Internet has its nickname, "word-of-mouse," delivered or enhanced by the social network channels of the Internet (Zilber, 1991). The formula for success requires an optimal mix of some free web content such as video clips, interactive games, blogs, e-books, and brandable software. This mix should offer valuable information that is ground-breaking, hilarious, or engages a celebrity, a network of people who actually light the fire, and has links that make the content very easy to share (Scott, 2008).

Steve Jurvetson, the venture capitalist behind Hotmail, originally coined the term *viral marketing* to describe how the service grew (Gordon, 2000). Hotmail started its business with viral marketing campaign in 1996: free e-mail. Hotmail users were converted to brand advocates by including a small message in every single e-mail they sent using the service, on the bottom of the note. The message said, "To get your FREE e-mail account go to Hotmail.com." (Thorne, 2008) The magic of this contagious marketing is that the product is the message carrier. The more you use Hotmail, the more you spread the virus (Gordon, 2000). The e-mail service spent only $500 000 on marketing to garner 12 million subscribers in a year and a half.

The form of viral marketing does not need to be sent as an e-mail. It can be as advergames, images, video clips, or text messages. YouTube, Facebook, MySpace, Twitter, and so on, are all good examples of venues for viral marketing. Videos are posted for all to see and anyone can forward a particularly liked clip to his or her own social network. Thus, viral marketing on the Internet is said to be the single most empowering and effective communication tool today (Scott, 2007). Undeniably, more and more marketers use this marketing magic (Thorne, 2008). However, a lot of marketers fail in getting their messages to be viral. The reason is that they do not understand fully how to create the content of viral marketing on the web, but instead still follow the rules of traditional marketing communication. Unlike advertising in traditional media which is expensive, the best word-of-mouse campaign inexpensively delivers content on the web in order to promote a marketer's organization and its products (Scott, 2008). The content should be directly connected to the company's products, services, and/or ideas. Viral marketing on the Internet is not a gimmick. Its success comes from self-replicating web content that people voluntarily share with their friends, colleagues, and family members.

The reasons that marketers increasingly include viral marketing as part of their integrated marketing communications are the power and cost-effectiveness that it offers. Viral marketing on the Internet is not a traditional top-down marketer-to-consumer communication, but instead takes a nonintrusive bottom-up approach, where consumers are in control of whether to interact with the communication message (Kirby, 2006). The user-driven process can ultimately help the marketer to achieve exponential brand endorsement by influencers and consumers. The cost-effectiveness of viral marketing on the Internet means that it provides accountability when tracked, thus measuring and proving ROI, and can even provide an increasing ROI (Kirby, 2006). In addition, it has no fixed cutoff point as long as the message remains contagious on the web. In conclusion, viral marketing has become an integral element in the overall marketing communication mix.

See also *e-commerce and Internet marketing; integrated marketing communication; international product diffusion; marketing functions on the Internet; marketing strategy implementation*

Bibliography

Gordon, S. (2000) *Unleashing the Ideavirus*, Do You Zoom, Inc, Dobbs Ferry.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Kirby, J. (2006) Viral marketing, in *Connected Marketing: The Viral, Buzz and Word of Mouth Revolution* (eds J. Kirby and P. Marsden), Butterworth-Heinemann, Burlington, pp. 87–106.

Scott, D.M. (2007) *The New Rules of Marketing and PR*, John Wiley & Sons, Inc, Hoboken.

Scott, D.M. (2008) The new rules of viral marketing: how word-of-mouse spreads your ideas for free, http://www. davidmeermanscott. com/documents/Viral_Marketing.pdf.

Thorne, L. (2008) *Word of Mouth Advertising Online and Off*, Atlantic Publishing Group, Inc, Ocala.

Wilson, R.F. (2005) The six simple principles of viral marketing. WebMarketing Today, February, http://www.wilsonweb.com/wmt5/viral-principles.htm.

Zilber, J. (1991) Of Mice and Menaces. MacUser, January.

opportunity identification

Henry Robben

Introduction

What is opportunity identification? Some authors would simply state that it is about getting ideas for new products (NPs) (Lehmann and Winer, 2005). In fact, it resides in the first phase of the process of NP or service development (SD), which consists of several stages or phases, varying in numbers depending on the source selected. It is strategic in nature, and "is the most difficult to describe or define" (Crawford and Di Benedetto, 2006, p. 28). As such, it is qualitative, informal, and approximate in nature rather than quantitative, formal, and precise (Kim and Wilemon, 2002). Most companies favor opportunity identification in advance of concept generation so as to demarcate the arena for the rest of the NP/SD process. However, others prefer opportunity identification to follow concept generation so as not to bias any directions in decision making (*see* CONCEPT TESTING).

Opportunity identification belongs to the so-called "fuzzy front end" (FFE) of the NP/SD process. Managers may find the FFE the weakest link in their innovation processes (*see* FRONT END OF INNOVATION). This first phase receives input from several activities, such as ongoing marketing and corporate planning and special opportunity analysis (Crawford and Di Benedetto, 2006). The result should be the identification of a number of opportunities through a creative process. Despite its recognized importance, little research has been devoted to the first phase in innovation management. Therefore, this process is not fully clear, as typically it is nonroutine, dynamic, and uncertain, but it concerns collecting business opportunities that could realistically be developed by an organization into successful products or services (Trott, 2008). Opportunities may arise from an underutilized resource (e.g., excess production capacity), a new resource (e.g., new skills or technologies), an external mandate (new regulations, a customer problem, a market threat or opportunity), or an internal mandate (e.g., Steve Jobs' ambition to "reinvent the phone" through the iPhone project) (Crawford and Di Benedetto, 2006). Opportunity identification, in fact, is the creative process by which an organization recognizes and assesses such opportunities. To be effective, it should be preceded and guided by a clearly formulated product strategy, for instance, a product innovation charter (Crawford and Di Benedetto, 2006), which basically sets out the strategy for opportunity identification. The product innovation charter provides focus on opportunity identification in the following ways:

- It provides a background: what are the reasons for developing a new offering? Which analyses have been done, and with what results?
- It necessitates focus: we need to describe and specify at least one technology and one market dimension.
- It specifies goals and objectives, for example, in terms of profit, growth, or market status.
- It identifies special requirements or guidelines, such as the innovativeness of the offering and the timing of market entry, possibly along with other guidelines, like an emphasis on avoiding a company's weakness.

Opportunities need not be restricted to a product or service concept. Kahn (2001) suggests new insights, new technology, or new customers can also be opportunities. As such, opportunity identification defines the hunting ground for further exploration. The process may differ depending on the size of the organization (e.g., large MNC versus a single branch), product type (e.g., FMCG, services, B2B versus B2C), and the strategic role of the product or service (e.g., breakthrough, incremental innovations).

Relevance

What companies are looking for are markets that are attractive – in terms of size in units and value, growth rate, profitability – with an opportunity to create a competitive position in one or more segments. The idea identified should therefore match the company's strengths and capabilities. The front-end activities result in the product concept (clearly aligned with customer needs), the product definition (explicit and stable), and

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

the project plan (priorities, resource plans, and project schedules).

Opportunity identification becomes especially relevant for industries such as pharmaceuticals. Investments are high, payoffs uncertain, and chances for differentiation are not easily obtained. Such industries experience the need to absorb much risk in the early phases of development (Bonabeau *et al.*, 2008). Proof of concept – an indication that the new product may be effective – is therefore needed for the idea to move on to the next phase in the new-product development process. The costs for moving an idea through the early phases of NP/SD are 1/5 to 1/50 of those for moving one through later stages.

For most organizations, it will not be possible to exploit all identified opportunities. That is why they need ongoing strategies covering NP/SD to systematically derive relevant criteria to judge the opportunities. Such criteria might include "do customers care, do we care, can we do it, can we stay ahead, or match the competition" (Griffin, 2000). Such criteria are necessary to assess the market or customer evaluation for new products. In essence, the process is the same for new-to-the-world products and close-to-home extensions. This part of the NP/SD process may also be referred to as "*a truth-seeking early stage, focusing on evaluating novel products' prospects and eliminating bad bets*" (Bonabeau *et al.*, 2008). As such, it tries to reduce uncertainty about a new product's or service's market promise quickly and effectively.

This reduction in uncertainty is a necessary step in a systematic process. In new-product decision making, managers can make several errors in judgment. They may ignore evidence challenging their assumption that a project would succeed for different reasons such as a shared belief that the new product will work. They also may terminate a project for lack of evidence that it could succeed. This may occur when a firm does not conduct adequate research or testing. Typically, many organizations focus on the later stages in development, when market introduction is near and management's attention is drawn by the costs of launching the product or service. This focus may in fact enhance the occurrence of both errors, as they neglect the importance of early testing.

Creating New Opportunities

The literature suggests many ways to generate new ideas that potentially create opportunities for companies. Trott 2008, for instance, identifies existing products, competitor's products and reverse engineering, technology, unexploited patents, customers and vendors, the sales force, senior and top management, brainstorming and other creative techniques, and individuals as sources for potential new business ideas (*see* CREATIVITY; BRAINSTORMING). Moenaert *et al.* (2009b) add to this a list of varying sources, like identifying underserved, overserved, or unserved market segments, listening at conferences, hiring consultants of competitors, and the analysis of garbage. Björk and Magnusson (2009) suggest that the better people are connected in a network, the higher the quality of ideas generated; interaction with other individuals increases the number of valuable new-product ideas.

An obvious source for new opportunities is the current customer, who may voice ideas for new features or benefits, or complain about an offering's current performance. Noncustomers or nonusers also form a source of new opportunities, as Nintendo's Wii game console shows: it was designed for people who were not interested in playing "traditional" computer games. Alternatively, a company may investigate a customer's value chain in terms of their consumption and production processes. Any inefficiencies or ineffectiveness identified may lead to a new business opportunity. Employees also form a source of new ideas, be they from R&D, manufacturing, sales, or intelligence units. Checking the competitive landscape for new opportunities may reveal competitors' actions or intentions, thus creating a sense of urgency for counteraction. For each of these sources, there are many creative ways to help come up with new opportunities, brainstorming being a classical example (*see* IDEA MANAGEMENT).

Assessing New Opportunities

In assessing new opportunities, we try to differentiate between potentially successful ideas and those that are not: "*The challenge is to separate the real opportunities from the illusions*" (Goold and Campbell, 1998:130). That is why

a new-product idea in fact faces two screens in the early phases. The first one controls whether an idea begins the phase or not. The second concerns whether the idea ends the FFE. In essence, one wants to manage two types of errors. A type I error occurs when a possible success is rejected; this may happen in rigid evaluation processes. A type II error occurs when failing to reject a possible failure; in this case, the evaluation process is weak. Ideally, we want to reduce both kinds of errors through our assessment of business opportunities.

Recent research suggests that managers evaluate business opportunities or strategic market options on the basis of four criteria (Moenaert, Robben, and Gouw, 2009a). These criteria are business opportunity, feasibility, competitiveness, and leverage. The first three criteria are must-haves, the latter one is a nice-to-have criterion.

Business opportunity. The first criterion involves the *business opportunity*, that is, the economic rent the manager expects to gain from the investment. In the eyes of the decision makers, business opportunity is the prize to aim for. This prize can contain a number of elements and have results for the short, medium, and long terms: increase in market share, increase in customer share, access to new markets or distribution channels, maintaining or increasing prices, increasing margins, increasing the efficiency of business processes, increasing the appropriability of revenues, increasing the economic value added, and so on. More precisely, each option represents a chance to create business or improve efficiency, or a combination of both.

Feasibility. A second dimension concerns *feasibility,* that is, the degree to which the proposed option is expected to be possible, in economic, technical, and organizational terms. Managers consider feasibility, together with opportunity, as the key criterion in judging strategic market options. Strategic projects often involve important changes in assets, know-how, and core processes. Such investments hence carry a high degree of organizational risk. Decision makers assess the feasibility of a strategic option on (i) its alignment with corporate strategy, (ii) the resource requirements, (iii) the flexibility, and (iv) the quality of the team that proposes the strategic option.

Competitiveness. The third dimension relates to the c*ompetitiveness* of the proposed strategic option, that is, the relative strength of the proposed new product as compared to the offerings of other firms in the industry. The competitiveness dimension consists of three key elements: (i) to what extent does the proposed option offer a competitive answer to the identified threats and opportunities; (ii) how big is the advantage over the competition the company seeks to develop – is it incremental or game changing? and (iii) how sustainable is the proposed option – how fast can a competitor come up with a similar solution?

Leverage. The leverage of a strategic option involves the expected likelihood of positive spillover effects. Such spillover effects may involve business leverage in other business units belonging to the same company (e.g., other business units producing similar products on the basis of a new platform). However, spillover effects may also result from future spin-offs that follow from this investment within the same business unit (e.g., a new-product platform enables a low-cost solution in the long run). Positive spillovers are "nice-to-have" factors that senior executives are likely to consider when making investment decisions. Leverage effects may be viewed as "outbound" synergies. It is crucial then, to adopt a realistic perspective on such leverage effects.

Moenaert, Robben, and Gouw (2009a) showed that ex-ante, feasibility and business opportunity dominated the decision-making process of managers, whereas ex-post the competitiveness of a new business idea appeared an important determinant of new-product success. Assessing strategic options is a process that requires inputs from different functions in the company. Analyses where financial data are lacking, or those where there is insufficient input from say, manufacturing, tend to be shallow. From experience we can tell that the quality of the assessment improves with the assessors having done

up-front data collection and analysis. Typically, senior management may play a stimulating and facilitating role in preparing the assessment process. Following this analysis, an investigation into IP, finance, and judicial matters can proceed.

The logic behind this assessment procedure is compelling. If the assessors determine that there is a business opportunity, that it is feasible, and that it creates a competitive position, the odds are that this particular option moves to the next phase. Options that do form a business opportunity, but are not feasible, or are not competitive, cannot add value to the company. An option that is also useful in other departments or markets or in the future, is a nice-to-have. It is not vital to new-product or -service success, but it may add a certain synergy among the company's NP/SD process.

There are many alternative evaluative mechanisms, for instance, the opportunity/threat matrix (Kotler and Keller, 2009). Two suggestions may be valuable in this respect. First, the criteria used for evaluating strategic options or business opportunities need to be unambiguous and formulated before the evaluation process starts. This formulation prevents unclarities from entering the evaluation process, making the assessment invalid. Second, following conventional wisdom that "all models are wrong, but some are useful," an assessment team should adapt existing models or build new ones to cater to the company's specific needs. Using standard models may ease the comparison with competitors, but to win new markets or customers, one needs to be different from or better than the competitor. Although the process of opportunity identification may seem like a generic component in the NP/SD process, companies should adjust it to suit their needs.

At the end of the evaluation of strategic options or opportunities, the assessing team has a systematically evaluated set of opportunities. This set can now act as an internal quote to a company's management to further pursue the evaluated opportunities in the NP/SD process, or as an external quote to venture capitalists, to acquire additional funds for further development. The next phase of the NP/SD process seeks to refine the results obtained, to refine the market definition, and to analyze promising opportunities into more detail, so as to come up with a product concept in terms of customer need description, form (physical representation for products or delivery mode for services), and the technology to create this form.

Conclusion

Most publications cite the phase of opportunity identification as a creative yet difficult-to-manage process. The issue at hand is not the generation of ideas. For this, we have many sources or techniques available. Once a pool of ideas (or opportunities, or strategic options) is formed, it is vital to choose the most promising opportunities among the ideas generated. To this end, it becomes important to use relevant and clear criteria to assess the viability of each idea. These criteria need to be defined early in the process, and be based on, for instance, a product innovation charter (*see* GROWTH STRATEGIES).

Bibliography

Björk, J. and Magnusson, M. (2009) Where do good innovation ideas come from? Exploring the influence of network connectivity on innovation idea quality. *Journal of Product Innovation Management*, **26** (6), 662–670.

Bonabeau, E., Bodick, N., and Armstrong, R.W. (2008) A more rational approach to new-product development. *Harvard Business Review*, **86** (3), 96–102.

Crawford, M.C. and Di Benedetto, A. (2006) *New Products Management*, 8th edn. McGraw-Hill, New York.

Goold, M. and Campbell, A. (1998) Desperately seeking synergy. *Harvard Business Review*, **76** (5), 131–143.

Griffin, A. (2000) Product decisions and marketing's role in new product development, in *Marketing: Best Practices* (eds M.R. Czinkota *et al.*), The Dryden Press, Orlando, pp. 248–289.

Kahn, K.B. (2001) *Product Planning Essentials*, Sage, Thousand Oaks.

Kim, J. and Wilemon, D. (2002) Focusing the fuzzy front-end in new product development. *R&D Management*, **32** (4), 269–279.

Kotler, P. and Keller, K.L. (2009) *Marketing Management*, 13th edn. Pearson Prentice Hall, Upper Saddle River.

Lehmann, D.R. and Winer, R.S. (2005) *Product Management*, 4th edn. McGraw-Hill, New York.

Moenaert, R., Robben, H., and Gouw, P. (2009a) *Marketing Strategy & Organization*, LannooCampus, Tielt.

Moenaert, R.K., Robben, H., Antioco, M. *et al*. (2009b) Strategic decision making and new product development. Paper presented at the 16th InternationalProduct Development Management Conference (EIASM). Twente, 7–9 June 2009.

Trott, P. (2008) *Innovation Management and New Product Development*, 4th edn. Prentice Hall Financial Times, Harlow.

product design

Robert W. Veryzer

Definition of Product Design

The phrase *product design* can refer to either the process of creating a product or the embodiment of a created item (a "product design"). Product design involves idea generation, concept development, and formulation that spans a number of considerations as well as knowledge or expertise domains – human needs, problem solving, synthesis of technology, ergonomics, manufacturing processes, environmental concerns/sustainability, market viability/acceptance, engineering, industrial design, marketing, and so on. The designing and designs of products to be manufactured and produced in quantity typically involves collaboration (*see* CROSS-FUNCTIONAL TEAM) among specialists from various disciplines such as engineering, industrial design, research and development scientists (*see* RESEARCH & DEVELOPMENT), market researchers, as well as others (Veryzer, 2005b). The aspects of design focused on by the various disciplines involved in producing (designing) a product design differ widely. For example, engineering design focuses more heavily on the function of a product and the means by which that functionality is delivered (e.g., mechanism) than does industrial design, which focuses more on product-use experience, user–product interfaces, ergonomics, functional enhancements, and concerns about aesthetics. Marketers are typically concerned with identifying and understanding market opportunities as well as collecting information from or about likely customers or users that will help shape a product's design (*see* LEAD USERS; VOICE OF THE CUSTOMER). They also focus on developing a strategy or program for ensuring a product's commercial success (*see* GROWTH STRATEGIES; NEW-PRODUCT FORECASTING), and the decisions that this entails frequently affect a product design in the form of "constraints" (e.g., size, weight, colors, desired product positioning, and price points). Although each of these disciplines may contribute to the overall product-design effort, they have disparate approaches and focus on different – yet often interrelated – product-design concerns (Veryzer, 2005b).

Translation of Needs into Product Design

Product design fundamentally arises from a problem or difficulty to be solved or a need or desire to be met. Although the front end of a product-design effort encompasses OPPORTUNITY IDENTIFICATION and analysis, as well as idea generation and selection, inspiration for a design, as well as the idea(s) for a design, may come from various sources (e.g., observations, nature, and toys), or in various ways – addressing a need/necessity, thoughtful contemplation, accidents, serendipity, and so on (*see* CREATIVITY). Regardless of the path by which product design is generated, a deep understanding of the problem or need (or desire) is crucial to both the product-design effort and the ultimate success of a product design. An understanding of the problem or need (or desire) to be addressed yields the core or key benefit(s) the product (design) is to provide for the person, group, or business for which it is being designed (Urban and Hauser, 1993; *see* VALUE PROPOSITION). Sometimes, discerning the dimensions of a problem or need (or desire) is relatively straightforward through observation, questioning, and trial; however, customers do not always know what they want, or are unable to sufficiently articulate what they actually need. Even though the underlying structure or dimensions of a need are not always readily identifiable or explicit, this does not discount the importance of gaining a deep understanding of both the problem/need and the consumer – through extended observation and in-depth study when necessary – as an important foundation for product design.

In interpreting and translating needs into designs, an understanding of the dual nature of many product designs both as a physical entity and a psychological/mental identity is essential for developing a complete and ultimately successful product design. This requires attention to a number of aspects of product design including product identity, usability, ergonomics, image, form, emotional appeal, and aesthetics, and these extend even to

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

hedonistic and psychological drives relating to self-image/self-esteem, desire, and sensuality (Veryzer, 2000). Woven throughout such considerations are currents of culture and fashion, along with more seemingly direct aspects of a product design like price and value – which may also have symbolic dimensions or meanings.

Along with a deep understanding of the problem, need, or desire, it is also necessary to have an appreciation of the elements of a design and the constraints relating to them. Product designs may be thought of in terms of the function(s) for which a product is being designed, requirements and constraints (specifications as to how the product must deliver the desired function – *see* PRODUCT SPECIFICATIONS), design parameters or vectors of potential design solutions that can deliver the desired functionality while satisfying the requirements and constraints, and constituents or the parts and assemblies (components) that are to make up the product design. Various methods such as house of quality or quality function deployment (QFD) may be used to structure the task of translating customer demands (needs) into design targets (*see* QUALITY FUNCTION DEPLOYMENT (QFD); CONCEPT SELECTION MATRIX).

Design Composition

In terms of specifying the product design, several dimensions should be considered. The first is the function of the product and how that functionality is to be safely delivered; the second is constituent or essential parts; and the third is positioning, inclusive of (product) identity, emotional appeal, and differentiation from other product offerings.

A product's configuration or *architecture* as it is sometimes termed, which addresses delivery of functionality(ies) by determining the necessary components and their layout, begins to emerge during the concept-development phase in sketches, function and component diagrams, as well as in early prototypes (Ulrich and Eppinger, 2008, p. 167; *see* PROTOTYPE). Technology (*see* TECHNOLOGY S-CURVE) may, in part, determine this, as does the degree to which the product or product category is established or has conventions. Product architecture may be thought of as a continuum ranging from completely integrated (intertwined parts) to individual, easily separated parts. In designs employing an "integral" architecture, functional elements of the product are highly interrelated and overlapping, and thus the boundaries between parts, components, and subassemblies are often difficult to identify or separate (Ulrich and Eppinger, 2008). In design executions employing a "modular" architecture (*see* PRODUCT MODULARITY), functional elements (parts and mechanisms) implement one or very few of the functions the product performs. The independence or "separability" of the physical elements of a product (design) usually makes it relatively easy to effect changes with respect to one aspect (part or function of a product design) without affecting other elements that make up a product. Product architecture can have significant implications for a product design on various dimensions. For example, for a particular instance, a highly integrated approach to product design may result in reduced assembly and manufacturing costs. Likewise, a modular approach to a particular product design might be used to provide cost and manufacturing efficiencies by using a common product platform (*see* PRODUCT PLATFORMS) as a basis for multiple product offerings. Product-design configuration has implications for the consumer or user as well as manufacturers. For example, a modular approach may lend itself to being easily upgraded, add-ons, adaptations or customizing, servicing, and even reuse of components.

The constituents vary according to what is suitable and required for a given product category (e.g., materials, fasteners, system components, switches, software, and finishes); these are virtually unlimited in terms of the possibilities and the different ways in which they may be configured. With the advances in electronic and digital technologies, there is greater latitude in configuring a product (design) than ever before. Beyond the direct concerns relating to the elements of a product design, the composition and configuration should take into consideration factors such as costs, manufacturability, available and standard components, available

materials and resources, manufacturing facilities, and manufacturing–product environmental impact (sustainable design).

A product design assumes an identity or "positioning" in the minds of those who encounter it (*see* PRODUCT POSITIONING). In other words, by design or default, people perceive a product design as conveying a meaning (e.g., "new," "efficient," "durable," "cheap," "fast"). There are a number of factors that may contribute to how a product is perceived and judged. These include aspects of a product design such as aesthetics, usability, performance, price (due both to customer resources as well as price–quality inferences), environmental impact, familiarity of the product form (as well as the product category), emotion(s) evoked, the degree of differentiation from competitors' product offerings, the novelty of the (product design) embodiment, and the degree to which the product is innovative or new. The task of creating a design identity or product position can be further complicated by dynamic factors such as fluctuating technologies, globalization (*see* GLOBAL PRODUCT DEVELOPMENT), cultures, trends, fashions, and images, which play a considerable role redefining products and product categories over time.

Product Design Strategy

Pursuit of marketplace objectives (profitability, market share, solving a difficult or previously unsolved problem, etc.) beyond the pure function(ing) of a product can play a substantial role in shaping a design (*see* COMPETITIVE ADVANTAGE). There are several fundamental product-design strategies that may be pursued in creating a product design (Veryzer, 2005a).

- *Cost/price advantage strategy*: emphasis on design so as to minimize component and production costs, thus affording a strategy of price domination.
- *Design eminence strategy*: establishing an identifiable and unique product image using, primarily, ergonomics, innovative concepts, and aesthetics so as to significantly elevate the product offering above the competition.
- *Feature leadership strategy*: pursuing new solutions and pushing the frontier in applying or developing technology (*see* DIFFUSION OF INNOVATION; INNOVATION METRICS; RADICAL INNOVATION).
- *Specific appropriateness (or concentration) strategy*: focusing on design that delivers the optimal product for a narrow product space (e.g., product category or the requirements of a particular type of user).
- *Desirable alternative strategy*: executing design consistent with an established product category or category leader but with sufficient variation to offer an attractive alternative choice.

In effect, each of the strategies involves or may result in a "position" for a product – for example, low cost, performance leader, most beautiful, preferred alternative, leading edge, and so on, and the position created by the design (composition) plays a key role in the marketing and promotion of the product. These strategies are applicable across design situations (*see* PRODUCT LIFE CYCLE) involving existing product bases (renovative and adaptive design), or new product formulations (e.g., evolutionary and discontinuous design).

Considerations for Product Design (Process)

There are a number of important considerations for product design and the product-design process, some of which are discussed in the following.

Expertise required.

- Various kinds of expertise are needed to delineate a product design and reconcile the inherent trade-offs (*see* INTEGRATED PRODUCT DEVELOPMENT). These range across such disciplines as engineering, industrial design, marketing research/marketing, and others; and encompass areas involving customer/market perspective, understanding culture(s), understanding user needs, opportunity analysis and product/market viability studies, product definition, concept generation, concept sketches, domain-specific principles, engineering (e.g., mechanical and software engineering), materials evaluation,

prototype development, product-use experience, aesthetics, interface design, human factors/ergonomics, emotional appeal, safety assessment, testing with customers, market confirmation, and manufacturing processes.

- The roles (market research, engineering, industrial design, marketing, etc.) may sometimes become blurred either because of various design team members being aware of the importance of considerations usually handled by someone from another discipline (this is usually constructive and beneficial to the design process and the product-design outcome), or when a (design) team lacks the expertise of one or more contributing disciplines (e.g., this has been observed in discontinuous innovation product-design settings where the process is often heavily R&D driven).
- There are numerous specialized product-design professionals who may be involved in the design process; they may be specific to particular design forms, industries, and product categories (e.g., automotive design (exterior design, interior design), graphic design, computer graphics design, appliance design, and toy design).

Design process.

- The product-design process may vary depending on the specific design and marketing goals; however, typically, the process involves (i) opportunity definition (with a deep understanding of the problem, need, or desire motivating the design effort); (ii) concept (solution) generation/ideation (*see* BRAINSTORMING), exploration, and development; and (iii) selection of design solution (*see* IDEA MANAGEMENT) and refinement (involves constructing product representations and prototypes; *see* PROTOTYPE; PRODUCT TESTING). Concept-confirmation research with target customers is usually conducted at various points throughout the process (*see* CONCEPT TESTING).
- The product-design process is usually iterative, subject to periodic evaluation and refinement of the emerging product design (*see* THE STAGE-GATE IDEA TO LAUNCH SYSTEM).

Market viability.

- Marketing activities directly related to product design tend to progress from market/opportunity definition to customer research (e.g., focus groups, and perceptual maps), to concept testing and product-design validation (with customers/users), test marketing, and eventually on to activities less directly related to the product-design effort such as product launch planning (e.g., determining appropriate promotion and advertising programs; *see* LAUNCH STRATEGIES).
- There are a number of marketing-related product-design goals. These involve design for speed to market, design to lower price (e.g., reduce costs, facilitate ease of manufacture, and use of common or easily adaptable platforms), design for identity (differentiation, positioning, brand, and corporate identity), universal design (to make a product design appropriate and usable for the broadest possible range of consumers), social-responsibility design (a desirable objective that is increasingly becoming an attribute sought by consumers), and marketing strategy compatibility (e.g., compact design so as to facilitate easy and low-cost distribution).

Bibliography

Ulrich, K.T. and Eppinger, S.D. (2008) *Product Design and Development*, The McGraw-Hill Companies, Inc., New York.

Urban, G.L. and Hauser, J.R. (1993) *Design and Marketing of New Products*, Prentice Hall, Englewood Cliffs.

Veryzer, R.W. (2000) Design and consumer research. *Design Management Journal: Academic Review*, **1**, 64–73.

Veryzer, R.W. (2005a) Enhancing new product development success through industrial design strategy, in *The PDMA Handbook of New Product Development*, 2nd edn (ed. K.B. Kahn), John Wiley & Sons, Inc., Hoboken, pp. 378–388.

Veryzer, R.W. (2005b) The roles of marketing and industrial design in discontinuous new product development. *Journal of Product Innovation Management*, **22** (1), 22–41. Special Issue – "Marketing Meets Design: Core Necessities for Successful New Product Development".

radical innovation

Gina Colarelli O'Connor

Definition

Radical innovation (RI) has been defined previously as innovations that offer either new-to-the-world performance features, or significant improvement (5–10 times) in known features or significant reductions (e.g., 50%) in cost, such that new application domains become possible (Leifer *et al.*, 2000; O'Connor, 1998). Technology, either newly discovered or in unique combinations, typically enables the innovation, and is recognized increasingly as the basis of the competitive advantage (Garcia and Calantone, 2002; Hill and Roethaermel, 2003; Lynn, Morone, and Paulson, 1996). RIs transform existing markets or industries, or create new ones. By this definition, computerized tomography (CT) and magnetic resonance imaging (MRI) were RIs in the field of diagnostic imaging, whereas the subsequent incremental and generational improvements in these technologies were not. The first personal computer was an RI, as it defined a new industry. Likewise, pagers and cellular telephones were RIs when first introduced.

RI is accompanied by high levels of uncertainty because, to achieve these performance leaps, firms must expand to incorporate or create emergent customer markets and/or new technology competencies. Each requires new knowledge or departure from existing skills or familiar practices and relationships (Christensen, 1997; Hill and Rothaermel, 2003; Levinthal and March, 1993). This not only triggers market and technical uncertainties but, also resource and organizational uncertainties (Leifer *et al.*, 2000). The eventual commercial success of RI opportunity is also highly uncertain. This, coupled with the enormous investment requirements to build new technical or market competencies given the novelty of the innovation for the firm, makes managing RI a highly risky endeavor, but one that promises the corporate renewal sought by large established firms.

Garcia and Calantone (2002) have criticized scholars for muddying the definitional waters regarding innovation. They call for a three-level typology of innovation, rather than the simplistic radical-incremental dichotomy that is so often used (*see* INNOVATION TYPOLOGIES). In fact, a survey of terms used in the industry reflects this three-level approach. Table 1 presents a variety of vocabularies used by firms today to reflect their experiences with innovation: most consider three different levels.

Garcia and Calantone (2002) develop their typology of innovativeness on the basis of micro (firm) and macro (market or industry) levels of impact as well as the presence of technical and market discontinuity, and define *radical innovation* as an extremely rare class of innovations that cause disruption at both the firm (micro) and the industry/market (macro) level, and that require new market creation and novel technology discovery. Their second category, *really new innovation*, which they describe as much more common, exhibits macro level discontinuity on *either* the market or technical dimension and *either* market, technical, or both types of discontinuity at the micro level. In whatever manner one might choose to classify these, what is clear is that both "really new" and "radical" innovations share a characteristic that incremental innovation does not – that of high levels of uncertainty on multiple dimensions. Incremental innovation is based on technology improvements along predictable trajectories, and is targeted toward a market that is familiar to the company. It relies on speed of execution and customer intimacy. But by definition, the other two categories require organizations to move into unchartered territory, where reliance on experience, current knowledge assets, and loyal customers are not an advantage. Radical and really new innovations require situation-specific knowledge and willingness to evolve business models to reflect new markets that are created. Both types may cause the firm to face issues not just of technical and market uncertainty, but organizational and resource uncertainty as well. The dynamics of competition change as well.

Companies do experience "really new" innovation much more frequently than RI, but the management challenges are of like kind, owing to the firm's inability to rely completely on current technology competencies and current customers,

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Terminology used by firms to describe levels of innovation.

Terminology Used	*Companies*
Horizon 1, 2, 3	IBM, 3M
Incremental, platform, breakthrough/radical	Gerber Scientific
Today, tomorrow, beyond	Corning
Incremental, major improvements, step-outs	Sealed Air
Incremental, substantial, transformational	J&J Consumer Products
Business unit projects, CEO projects, advanced technology projects	GE
Aligned, white spaces, grey spaces	Air Products
Quadrant 1,2,3,4 (alignment vs. time horizon)	Kodak
Incremental, platform, scope change (Charles River)	Mead-Westvaco
Incremental, gamechangers	Shell Chemicals

and the need to learn, redirect and reformulate given these uncertainties. The difference between "really new" and "radical," is one of degree (i.e., uncertainties may exist on more dimensions for "radical" than for "really new"), and so, for purposes of this article, they are taken together, and considered as "major" innovations. The difference between these two forms of "major" innovation and the incremental innovation is one of kind: managing for innovation in which uncertainty levels are high requires approaches that differ from those used when previous experience is of value and uncertainty levels are lower.

Another perspective used to define RI is the nature of its impact on the world of competition. Utterback, (1994) defines RI as that which introduces discontinuity into industry dynamics, leaving incumbent firms' investments in technical skill, knowledge, designs, plant, and equipment no longer providing competitive advantage. The effect is to dislocate incumbent firms from their leadership positions and completely restructure industries on the basis of innovations offered by outsiders. IBM's commercial success with the electric typewriter unseated Olivetti and other industry players in the manual typewriter industry. In the lighting industry, oil lamps displaced the use of candles, but were themselves displaced by gas lamps, then electric lamps, fluorescent lamps, and now halogen lamps. Each wave of change unseated the industry leader, with the exception of GE with their successful adoption of fluorescent technology.

Chandy and Tellis, (1998) note that most definitions of RI share the commonality in (i) the extent to which the product incorporates a new technology and (ii) the extent to which it brings dramatic increases in value to the marketplace relative to current solutions. An RI is an offering that is high on both the technology and market dimension. A market breakthrough may bring dramatic increase in benefit to the market without novel technology. A technological breakthrough may use substantially novel technology without change in the perceived value in the marketplace.

Sood and Tellis, (2005) question the definitions that rely on market impact and instead focus on novelty of the technology. They suggest that defining an innovation in terms of its effects rather than its attributes is circular, and revert to defining platform innovation, instead, as the emergence of a new technology based on scientific principles that are distinctly different from those of existing technologies. While their arguments regarding tautology are important, one must consider the level of analysis one is examining. Perceived, if not real, market impact is a critical component of the definition for those studying management practices associated with RI, particularly where market discontinuities are expected to occur.

Christensen (1997) notes that RIs may be either disruptive or sustaining. In either situation, they are defined as offering new-to-the-world performance features or dramatic improvements in a key feature of interest.

All of these perspectives share the following common elements:

1. *Immense learning and novelty for the firm leading the commercialization effort*: This requires leveraging current competencies but, more importantly, developing new ones, on both the market and technical dimensions. It requires different management systems that are aligned with an innovation objective and a recognition of uncertainty and ambiguity. This also impacts their strategic focus, and may force decisions about the company's strategic intent, that is, its investments for the health of the organization in the distant future.
2. *Disruption of competitive dynamics*: Established competitors may not be the key competitors; rather, firms and technological approaches that are external to the firm's current competitive set may become prominent figures.
3. *New market formulation*: Agents that never previously participated may be required to join the value network, the ecosystem of participants necessary to bring the innovation to fruition. This may require external incentives for them, given the uncertainty of the market opportunity.
4. *Changes in consumer behavior*: Most RIs require the adoption of new usage techniques, and therefore an investment in learning, on the part of buyers.

Radical Innovation and the Field of Marketing

The phenomenon of RI is rife with marketing issues. It is therefore important for marketing scholars to engage in research to deepen their understanding. Points of intersection between RI and marketing may include seeking answers to the following questions

- How do firms recognize opportunities for breakthrough innovation (BI), given the lack of an identified market a priori?
- What approaches work best for learning about markets that do not exist?
- What marketing and entrepreneurial interfirm practices are required for successfully commercializing RIs. What practices diminish the barriers to such success?
- What practices can firms and other institutions take to create new markets?
- How will marketing professionals span the boundary generated by the increased need for interfirm partnerships in finding breakthrough opportunities and commercializing them?
- What can be learnt about diffusion of RIs and about what impacts market take-up rates? Price has been examined, but what other phenomena influence the take-up timing and rate?
- What are the competitive dynamics for RI in comparison to conventional fast-paced development cycles?
- What is the role of an organization's market power in an industry in suppressing and/or elevating RIs?
- What influences consumer behavior in the adoption of RI?

Industry Level Analyses of Radical Innovation

At the industry level of analysis, work has centered on understanding technological evolution and the role of RI, irrespective of whether an economic advantage exists for pioneers of RI, and the impact of specific competitive dynamics (e.g., price) on stimulating markets to adopt RIs.

Technological evolution and RI. The literature describes a technology as evolving through an introductory period of slow growth as it becomes better understood and more researchers become aware of it. This is followed by a period of fast growth as the scientific understanding passes a threshold and a dominant standard emerges around which numerous product platforms can be defined (*see* PRODUCT PLATFORMS). Finally, after a period of improvement the technology reaches a mature stage when a ceiling or plateau in improvement is reached, and the focus of innovation shifts from product to process innovation (Utterback, 1994).

Graphically, this phenomenon is represented by an S-curve.

Recent works by Sood and Tellis, (2005) find that this widely held theory is not confirmed in a number of cases they empirically examine. Rather, they find that distinctly novel or "platform" technologies evolve through an irregular step function with long periods of no growth in performance or technological improvement, followed by performance jumps. Thus, multiple S-curves may be a more appropriate explanation of technology evolution than one singular curve (*see* TECHNOLOGY S-CURVE). Additionally, contrary to previous research (Christensen, 1997), Sood and Tellis find that the performance curves of two rival technologies may experience more than one point of intersection, and rivals may successfully attack from below a leader's performance trajectory. Thus, the final status of each technology cannot be determined from the direction of the attack or timing of introduction. Novel technologies are introduced as much from new entrants as from large incumbents. The rate of technological change and the number of new technologies increase over time within any new technology platform. Additionally, each new technology introduces a set of secondary dimensions as a new basis of competition. Practically, these results suggest that premature abandonment of investments in novel technology may occur for firms persuaded by the S-curve theory.

Pioneering and radical innovation. While much of the work in marketing regarding pioneer advantage is based on products that may be new but not necessarily radical, some attention has been paid to pioneering with advanced technology or "really new" products (Min, Kalwani, and Robinson, 2006; Bayus, Jain, and Rao, 1997; Klepper and Simons, 1997). The results of these studies, while conclusive individually, lead to contradictory results collectively (*see also* FIRST-MOVER ADVANTAGE). The question of enduring market superiority for RIs, therefore, cannot be predicted solely by order of entry.

Min, Kalwani, and Robinson (2006) posit that extremely high levels of market and technical uncertainty faced by pioneers of really new products (in contrast to pioneers of incremental innovations) make survival risky. They demonstrate that 23% of market pioneers of really new industrial goods survived at least 12 years, compared with 61% of market pioneers of incremental innovations. Market pioneers creating really new products must also create new markets, face resistance, and manage technical uncertainty not experienced by market leaders for incremental innovation. Fast follower survival rates did not differ significantly, however, for really new versus incremental innovations, at 38 and 39%, respectively. Early followers benefit from the prior efforts of pioneers in building primary demand for really new products, learn from the pioneers' efforts to help customers articulate latent needs, and to leapfrog the pioneers' technical solutions.

Bayus, Jain, and Rao (1997) offer a more complex story in their study of the personal digital assistant market. An historical examination of market entrants and their differing product design decisions lead the authors to conclude that pioneers and early entrants rushed to market with a klugey device hoping to capture a mass market at the outset. Later entrants, however, focused on niche applications, introducing devices that were fully functional and specifically designed to meet the needs of the specialized markets they had targeted. Initial entrants' performance was dwarfed by that of later entrants. While the authors suggest that early entrants offered "too little, too early" one wonders what would have happened if they had not stimulated market interest with those initial devices.

A final study offers insight into the link between major innovations and early market entry. Klepper and Simons (1997) suggest that long-term survivors of industry shakeouts are those who entered early and who invested consistently in R&D such that they offered first and subsequent innovations, in both product and process. They find evidence of increasing returns to R&D, indicating that those who invest early in innovation and continue to do so subsequent to their initial entry dominate competitors over time. This conclusion, in conjunction with Tellis and Golder's (1996) observation that those who endure as market leaders exhibit vision, commitment, persistence, and continued investments in innovation brings some reality to the

quest for the silver bullet of the role of first entrant that researchers continuously seek.

Pricing and radical innovation. The scholarly work on causes of sales takeoff for really new products emphasizes new-firm entry or declining prices as key determining factors (*see also* TAKEOFF). Bayus, Kang, and Agarwal (2007) examine whether price declines occur because of monopoly profit maximization behavior in the early stages of a new market or because of competitive pressures unleashed by firm entry. They study 30 successful consumer and industrial product innovations deemed significant by a panel of experts in the field over a 125+ year period of time, and conclude that, in contrast to other studies, price declines are not direct causes of industry takeoff, determined by strategic choices made by market pioneers, but instead that increases in rates of new-firm entry into the market, spurred by the prospect of supranormal returns, drive down prices. Both prices and heightened rates of innovation due to increased numbers of competitors together stimulate market takeoff.

Firm Level

It is widely understood that large established companies are capable of creating and commercializing new products that offer incremental benefits to the market, but that they struggle in their attempts to develop new-to- the-world, RIs that confer whole new business and rejuvenation opportunities (Burgelman, 1983; Christensen, 1997; Dougherty, 1995; Fast, 1979; Leifer *et al.*, 2000). Empirical evidence across a variety of industries is mounting, however, linking RI to superior economic performance (Cho and Pucik, 2005; Sorescu and Spanjol, 2008), so the need to overcome these challenges and develop capabilities for radical and other forms of major innovation appears to be a critical issue for established firms and continues to command scholarly attention.

While RI is an important path to generative growth that established organizations might employ, evidence suggests that forces within such organizations can impede its success (Dougherty, 1995; Leifer *et al.*, 2000). Organizations rely on established routines (Nelson and Winter, 1982): procedures that promote efficiency and automation over creativity and flexibility. They grow by gaining efficiencies of scale and scope in core competency areas that ultimately turn into core rigidities or core incompetencies (Dougherty, 1995). Reliance on established routines, which is key to the established firm's ability to earn profits in known, well-established markets, reduces an organization's ability to cope with uncertainty in any manner other than to ignore it or elect not to engage in opportunities that introduce it. This precludes most opportunities that could result in RI, since each breakthrough, by definition, is unique, and situation-specific knowledge must necessarily be developed. For RIs, relying on past experience is not necessarily an aid to current opportunity. Technical uncertainties abound for most breakthroughs, but so do high levels of uncertainty and discontinuity in markets, resources, and organizational changes required to successfully commercialize them (Chandy and Tellis, 1998; Leifer *et al.*, 2000; Rice *et al.*, 2008).

Empirical evidence points to the fact that new venture groups or other incubator-type models in large organizations do not last long enough for their portfolios of innovations or the development of new innovation competencies to have an impact on the organization (Fast, 1979; O'Connor and Maslyn, 2002). Other scholars argue that organizations can develop appropriate management systems for RI but simply have not (Hill and Rothaermel, 2003; Leifer *et al.*, 2000; O'Connor *et al.*, 2008).

Sustained commitment to building an RI capability is a perennial challenge for firms, since the twin objectives of exploring (developing new lines of business through managing for breakthroughs) and exploiting compete for resources, and involve drastically different management systems. While ambidexterity is perceived to be a virtue (Tushman), the processes associated with exploitation appear to have the upper hand in managerial decision making, primarily because the connection of managerial decision to observed consequence is so much clearer than in the case of exploration activities (March, 1991). Thus, firms invest in improving exploitation practices at the expense of ignoring the development of good practice for exploration,

or new business creation. Yet, profits gained through innovative activities are termed *Schumpeterian'* or *entrepreneurial rents* because they are the rewards to firms who are prepared to act in the face of ex ante uncertainty (McGrath and Nerkar, 2004). The motivation to engage in potential RI s is thus always present within firms, for financial as well as employee motivational reasons.

From People to Processes to Management Systems

Reliance on individuals. The term *intrapreneuring* was popularized in the mid 1980s (Pinchot, 1985) to reflect the view that major innovation happens when corporate entrepreneurs break rules and create major innovations in spite of corporate resistance (*see also* ORGANIZING FOR INNOVATION; LEADERSHIP ROLES IN PRODUCT DEVELOPMENT). Once the innovation begins to exhibit success in the marketplace, leadership "retroactively rationalizes" their strategies to incorporate the new initiative (Burgelman, 1983).

Much of the RI literature at the firm level highlights the critical role of the single-minded, impassioned champion (Pinchot, 1985). Recent research, however, is documenting management practices that can help, rather than hinder the management of RI, thus reducing sole reliance on individual intrapreneurs. From a management-theory perspective, the central tenet of dynamic capability theory holds that firms can evolve processes that enable them to develop, change, and rejuvenate themselves. While persistent, skilled, visionary champions are critical to RI, organizations will never maximize their resources and advantages if they do not move from relying on champions to a dynamic capability that includes them as one element of a system.

New processes and structures. A number of organizational experiments, such as corporate venture divisions and external incubators (Xerox Parc), have been invoked as approaches to liberating BIs from the dominant operational excellence orientation in the firm, but these have suffered fairly poor results. New venture divisions, designed to incubate breakthrough or newstream opportunities until they can compete for resources and attention with mainstream product lines do not, in general, sustain themselves or receive organizational support for more than four or five years, since the fledgling businesses do not achieve large-scale growth quickly enough and managerial patience wears thin Fast, 1979). The famed Xerox Palo Alto Research Center (PARC), designed as an ongoing skunk-works group to develop such a portfolio, failed as a source of growth and renewal for its mother company, as its innovations were adopted and commercialized by others but rejected by its own sponsoring company. PARC was criticized for being too isolated from the company. Yet, integrating RIs into established businesses is a perennial problem, given the mismatch of potential business models, value propositions, customer usage patterns and market application, and the threat of cannibalization of current offerings that the receiving unit must face (Chandy and Tellis, 1998).

When BIs are governed as a collective, portfolio management processes become important. When BIs occur as a result of an individual champion managing independently, no such portfolio level oversight or thinking comes into play. Consideration of RI portfolios has led to advances in techniques for evaluating such portfolios and has resulted in the development of real options-oriented approaches for evaluating and monitoring the value of an entire portfolio, at the same time, evaluating each project within it (McGrath and Nerkar, 2004).

Management systems for major innovation. Companies have tried many approaches in recent history, but none of the models has resulted in a sustained successful BI capability. Dependence on either an individual champion or a lack of sophistication regarding appropriate processes and structures may be part of the problem, as evidenced in the literature to date.

A recent perspective that is emerging suggests that, while each of these is a contributing factor, it is the lack of viewing innovation as requiring a management system, in which all of these elements are considered simultaneously and managed in accordance with an innovation objective, that is the fundamental cause of such

poor performance of firms with respect to RI in the long run (O'Connor, 2008; O'Connor *et al.*, 2008). Treating BI as a temporary program, an appendage to the mainstream organization, or a one-off project that the company invests in on an ad hoc basis is risky. It makes innovation vulnerable to the whims and fancies of whoever is in charge, the current financial picture of the firm, or the condition of the capital markets outside the firm. The organization structure for managing a portfolio of RIs, along with appropriate processes, talent recruitment and development, governance and decision making, culture and leadership, and resource systems and metrics, together, comprise a management system for innovation that can help the firm maintain ambidexterity over time.

Project Level

The innovation literature questions the extent to which processes needed to manage an RI mirror those used for traditional new-product development projects, given the high degrees of uncertainty and risk for the former. Sethi and Iqbal (2008) find that the use of Stage-Gate approaches for RIs are counterproductive. A more learning-oriented approach is emerging, that recognizes the need for search in nonlinear pathways, iterative learning, experimental processes, and the development of new capability for each innovation (Bhardwaj, Camillus, and Hounshell, 2006; Cheng and Van de Ven, 1996; Danneels, 2002; Lynn, Morone, and Paulson, 1996; McGrath and MacMillan, 1995; Rice *et al.*, 2008; Sommer and Loch, 2004).

These results are critical for management practice, since search techniques resulting in learning but not in perceived project progress may be viewed by project review boards as failed attempts to achieve milestones. Additional work on project-level oversight is needed to enhance the applicability of the recent research progress on project management techniques to real-world settings. In terms of project-level oversight, Jansen *et al.* (2006) demonstrate that centralized decision making negatively impacts exploratory innovation, implying that senior-level managers may not be appropriately versed regarding how to monitor and engage with BI project teams at this juncture.

Future Directions

Rapid gains in understanding of RI, at all levels of analysis, are being made in the academic literature. The industry level of analysis has reached a level of sophistication with regard to the amount and richness of data that is being used to draw conclusions. This reliance on sources of data previously unaccessed has led to the toppling of a fair number of widely held principles that were based on case analyses and conventional wisdom.

At the firm level of analysis, the gains in dynamic capability theory, along with the invocation of the ambidexterity claim, or March's exploration/exploitation theory is leading to rapid developments in the literature that go beyond descriptions of single case studies and exhortations that firms must be ambidextrous to a serious discussion about how that can happen for the world of innovation.

Project-level analyses have additionally moved from the case-descriptive mode to the theoretical and prescriptive, indicating a fairly thorough understanding of the approaches to RI project management and oversight that may be more likely to lead to success, given the level of risk involved.

Bibliography

Bayus, B.L., Jain, S., and Rao, A.G. (1997) Too little, too early: introduction timing and new product performance in the personal digital assistant industry. *Journal of Marketing Research*, **34**, 54–63.

Bayus, B.L., Kang, W., and Agarwal, R. (2007) Creating growth in new markets: a simultaneous model of firm entry and price. *Journal of Product Innovation Management*, **24**, 139–155.

Bhardwaj, G., Camillus, J.C., and Hounshell, D.A. (2006) Continual corporate entrepreneurial search for long-term growth. *Management Science*, **52** (2), 248–261.

Burgelman, R.A. (1983) A process model of internal corporate venturing in the diversified major firm. *Administrative Science Quarterly*, **28**, 223–244.

Chandy, R.K. and Tellis, G.J. (1998) Organizing for radical product innovation: the overlooked role of willingness to cannibalize. *Journal of Marketing Research*, **35**, 474–487.

Cheng, Y.T. and Van de Ven, A.H. (1996) Learning the innovation journey: order out of chaos? *Organization Science*, **7** (6), 593–614.

Cho, H.-J. and Pucik, V. (2005) Relationship between innovativeness, quality, growth, profitability and market value. *Strategic Management Journal*, **26**, 555–575.

Christensen, C.M. (1997) *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*, Harvard Business School Press, Boston.

Danneels, E. (2002) The dynamics of product innovation and firm competences. *Strategic Management Journal*, **23**, 1095–1121.

Dougherty, D. (1995) Managing your core incompetencies for corporate venturing. *Entrepreneurship, Theory and Practice* (Spring), **19** (3), 113–135.

Fast, N.D. (1979) Key managerial factors in new venture departments. *Industrial Marketing Management*, **8**, 221–235.

Garcia, R. and Calantone, R. (2002) A critical look at technological innovation typology and innovativeness terminology: a literature review. *Journal of Product Innovation Management*, **19**, 110–132.

Hill, C.W.L. and Rothaermel, F.T. (2003) The performance of incumbent firms in the face of radical technological innovation. *Academy of Management Review*, **28**, 257–274.

Jansen, J.J.P., Van Den Bosch, F.A.J., and Volberda, H.W. (2006) Exploratory innovation, exploitative innovation, and performance: effects of organizational antecedents and environmental moderators. *Management Science*, **52** (11), 1661–1674.

Klepper, S. and Simons, K.L. (1997) "Technological extinctions of industrial firms" an inquiry into their nature and causes. *Industrial and Corporate Change*, **6** (2), 379–460.

Leifer, R., McDermott, C.M., O'Connor, G.C. *et al.* (2000) *Radical Innovation: How Mature Companies Can Outsmart Upstarts*, Harvard Business School Press, Boston.

Levinthal, D. and March, J.G. (1993) The myopia of learning. *Strategic Management Journal*, **14**, 95–112.

Lynn, G., Morone, J.G., and Paulson, A.S. (1996) Marketing and discontinuous innovation: the probe and learn process. *California Management Review*, **38** (3), 8–37.

March, J.G. (1991) Exploration and wxploitations in organizational learning. *Organization Science*, **2**, 71–87.

McGrath, R.G. and MacMillan, I.C. (1995) Discovery driven planning. *Harvard Business Review*, **73**, 44–54.

McGrath, R.G. and Nerkar, A. (2004) Real options reasoning and a new look at the R&D investment strategies of pharmaceutical firms. *Strategic Management Journal*, **25** (1), 1–21.

Min, S., Kalwani, M.U., and Robinson, W.T. (2006) Market pioneer and early follower survival risks: a contingency analysis of really new versus incrementally new product-markets. *Journal of Marketing*, **70**, 15–33.

Nelson, R.R. and Winter, S.G. (1982) *An Evolutionary Theory of Economic Change*, Belknap Press, Cambridge.

O'Connor, G.C. (1998) Market learning and radical innovation: a cross case comparison of eight radical innovation projects. *The Journal of Product Innovation Management* (Special issue on Really New Products), **15** (2), 151–166.

O'Connor, G.C. (2008) Major innovation as a dynamic capability: a systems approach. *Journal of Product Innovation Management*, **25** (4), 313–330.

O'Connor, G.C. and Maslyn, W.T. (2002) Nortel Network's Business Ventures Group: One Corporation's Take on Entrepreneurship, (revised October 2004), Harvard Business School Case #BAB057-PDF-ENG.

O'Connor, G.C., Leifer, R.L., Paulson, A.S., and Peters, L. (2008) *Grabbing Lightning: Building a Capability for Breakthrough Innovation*, Jossey Bass.

Pinchot, G. (1985) *Intrapreneuring: Why You Don't have to Leave the Company to Become an Entrepreneur*, Harper & Row, New York.

Rice, M.P., O'Connor, G.C., and Pierantozzi, R. (2008) Implementing a learning plan to counter project uncertainties. *Sloan Management Review*, **49** (2) (Winter), 54–62.

Sethi, R. and Iqbal, Z. (2008) Stage-gate controles, learning failure, and adverse effect on novel new products. *Journal of Marketing*, **72**, 118–134.

Sommer, S.C. and Loch, C.H. (2004) Selectionism and learning in projects with complexity and unforseeable uncertainty. *Management Science*, **50** (10), 1334–1347.

Sood, A. and Tellis, G.J. (2005) Technological evolution and radical innovation. *Journal of Marketing*, **69**, 152–168.

Sorescu, A.B. and Spanjol, J. (2008) Innovation's effect on firm value and risk: insights from consumer packaged goods. *Journal of Marketing*, **72** (2), 114–132.

Tellis, G.J. and Golder, P.N. (1996) First to market, first to fail? Real causes of enduring market leadership. *Sloan Management Review*, **37** (2) (Winter), 65–75.

Tushman, M.L. and O'Reilly, C.A. III (1996) Ambidextrous organizations: managing evolutionary and revolutionary change. *California Management Review*, **38**, 8–30.

Utterback, J. (1994) *Mastering the Dynamics of Innovation*, Harvard Business School Press.

accelerated product development

Fred Langerak

Introduction

Accelerated product development is a competitive strategy that seeks to compress the development cycle time of new products. However, there has been little theoretical advancement and empirical model testing regarding when cycle time reduction is appropriate, what factors accelerate product development, and how cycle time reduction affects project outcomes. Since cycle time reduction has become important for managing new-product development, this article defines, organizes, and integrates the literature on accelerated product development. Specifically, we argue that the antecedents that shorten development cycle time are well known, but that there are contradictions in empirical support regarding its outcomes in terms of development expense, product quality, and new-product success. We attribute these inconclusive results to contingency effects, methodological differences, trade-offs in cycle time reduction, the hidden costs of accelerated product development, and differences in new products' windows of market opportunity.

Definition

The popularity of accelerated product development is based on the contention that being fast can facilitate either first-mover or second-mover strategies (Kessler and Chakrabarti, 1996). The faster a firm can develop a new product, the greater the likelihood that it can be the first to market with a new product and reap pioneering advantages (*see* FIRST-MOVER ADVANTAGE). However, being first to market is not always a guarantee for higher sales levels, particularly when a firm's pioneering advantages are dependent upon the development cycle time of its followers. In other words, a fast imitation strategy on the part of its followers can reduce the firm's first-mover advantages. Moreover, the faster a follower can develop a new product, the more lead time it can put between itself and later movers. The implication is that both pioneers and fast followers should shorten their development cycle time to build a competitive advantage over later entrants (Chen, Reilly, and Lynn, 2005). Against this background, the literature defines product development cycle time as the time elapsed between the initial development, which includes conception and definition, and commercialization, namely, the introduction of the new product in the marketplace (Griffin, 1997; Kessler and Chakrabarti, 1996). Consistent with the notions of product development time, time to market, innovation time, total time, and lead time, this definition implies that development cycle time can be reduced by increasing the new product's development speed, innovation speed, or speed to market.

Literature Review

The results of consecutive studies of product development best practices by the Product Development and Management Association (PDMA) reveal that product development cycle times have decreased significantly over the past 15–20 years. During the same period, the academic and practitioner-oriented literature on accelerated product development has considerably broadened, and development cycle time has been written about extensively. However, accelerated product development is still one of the least studied subjects in new-product development.

Theoretical grounding. Studies on accelerated product development may be categorized into four streams of research. The first category encompasses grounded-theory approaches and small-sample studies conducted to uncover the drivers of product development cycle time (see Kessler and Chakrabarti, 1996 for a comprehensive review of this literature). The output of these efforts are numerous factors assumed to be associated with development cycle time, including project strategy features (e.g., product complexity, strategic intent, level of innovativeness, and technical difficulty), development process characteristics (e.g., formality, stages, and structure), and organizational characteristics (e.g., team use and assignment level, leadership style, size, and innovation level). Exploratory in nature, this line of research provides a solid

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

ground upon which theory can be developed (Griffin, 2002).

Antecedents of cycle time reduction. The second category synthesizes these exploratory findings to develop conceptual models and test the hypothesized influence(s) of accelerated product development on project strategy, and process and organizational characteristics on development time (a review of this literature is provided by Griffin (1997)). The empirical results relating to project strategy are unequivocal: newer, more complex, more technically challenging, and more innovative projects are typically associated with longer development cycle times. This suggests that firms striving for shorter development cycle times would need to undertake less complex, less innovative, and less technically demanding projects. While such a strategy may (indeed) reduce cycle time, it clearly carries long-run threats to marketplace success and long-term financial rewards (*see* INNOVATION METRICS; SUCCESS FACTORS FOR NEW-PRODUCT DEVELOPMENT).

The numerous studies on development process factors as antecedents of cycle time produce less clear results. Clear project (time) goals, partial or complete concurrent processing, acceleration of activities and tasks, detailed process planning, greater investment on both human and financial resources, and, at the fuzzy front end, increased rewards for R&D performance, lead user involvement, and a long-term orientation are associated with decreases in the development cycle time (Calantone and Di Benedetto, 2000). Development processes that use design for manufacturability tenets, new-product screening models, computer-aided design systems, and frequent product testing, and that display high levels of supplier involvement and greater number of customers involved with prototypes are linked to longer lead times (Griffin, 1997). These opposing effects suggest that firms making changes to their development process must closely monitor the impact of their interventions on both development effectiveness and cycle time efficiency, and recognize that a trade-off may be inevitable.

Other studies have identified a number of organizational actions that firms can take to reduce development cycle time, in particular improving the R&D–marketing–manufacturing interface by establishing cross-functional teams, and simplifying the organizational structure. Increased knowledge levels of team members, greater dedication on part of the project leader, and adoption of a more participatory leadership style have been found to be associated with shorter cycle times (Griffin, 2002). Thus, the empirical results on the antecedents of accelerated product development demonstrate that a considerable number of project strategy, process, and organizational factors are closely associated with product development cycle time. However, an overarching theory to the results across the aforementioned studies is yet to be articulated. The only exception is perhaps Gerwin and Barrowman's (2002) meta-analytic finding that overlap and interaction, tools and formal methods, and team leader influence work toward reducing development cycle time.

Outcomes of accelerated product development. The third stream of research comprises studies that investigate the outcomes of development cycle time reduction in terms of development costs, product quality, and project success. The studies in this category use multiple methodologies ranging from broad-based surveys and case analyses to simulations, and systematically test the hypothesized effect of cycle time reduction on the three sets of outcomes. The results of these studies do not provide unanimous evidence in favor of accelerated product development. With regard to the effect on development costs, some researchers have found negative correlations, while others have established that shortening development time lowers development costs (see Kessler and Chakrabarti, 1996 for a detailed review). The valence of the relationship between cycle time and product quality is also unclear. One team found that higher product quality is related to decreases in cycle time, while others document its association with increases in cycle time (Griffin, 2002). There is also little empirical support for the notion that reduced product development cycle is a key ingredient to project success. In a review of prior empirical studies, Chen, Reilly, and Lynn (2005) conclude that the literature has produced inconsistent, even conflicting, results on the relationship between development cycle time and project success.

In sum, while many view the realization of shorter cycle times as an important means to reducing development costs, improving quality, and attaining better project success, there seems little empirical support that substantiates this contention.

Contingency effects. The fourth category of research reflects on these divergent empirical results via its employment of (mainly) the survey methodology to examine the moderating effects of contextual factors, such as uncertainty, product innovativeness, new-product strategy, team improvisation, and customer participation. Mixed results have been documented in relation to market uncertainty, with some studies suggesting a weaker association between speed to market and project success, and others revealing higher correlations between these variables under conditions of uncertainty. There is, however, little dispute surrounding the finding that technological uncertainty has little, if any, effect on the speed–success relationship. With regard to project innovativeness, several studies suggest that innovativeness weakens the effect of product development cycle time on project success (Ali, 2000). Concerning strategy, findings imply that pioneers and fast followers should not use the same acceleration approaches, as the impact of the majority of these approaches on cycle time and profitability is moderated by the new-product strategy of the firm (Langerak and Hultink, 2005). Moreover, reduction of development cycle time is generally considered more essential to fast followers than for pioneers since it can help reduce the pioneer's lead time over later entrants. As far as improvisation is concerned, findings show that team improvisation in the context of a structured development process increases the likelihood of reduced cycle time and increased new-product profitability. Finally, research shows that customer participation as an information resource has a positive effect on speed to market when downstream customer connectivity is high, but no significant effect when it is low. In addition, customer participation as codeveloper undermines new-product speed to market when process interdependence is high. When interdependence is low, the effect of customer participation as codeveloper on speed to market is significant and positive.

Differences Related to Measurement Methods

The conflicting results of research on accelerated product development are commonly attributed to methodological differences in measurement approach and unit of analysis, and the relative importance of development cycle time reduction for new-product success (Chen, Reilly, and Lynn, 2005).

Measures. One source of inconsistency across studies stems from different assertions regarding the appropriate starting and end points of the product development cycle, and the consequent differences in cycle time assessment. Moreover, many studies suffer from a lack of rigor in data presentation, rendering the comparison of measures problematic in the face of different time frames (Griffin, 1997). Other discrepancies arise due to the variability in the conceptualization of cycle time and the diversity of approaches employed for its measurement. One approach, for instance, uses the actual elapsed time between the spark and the launch of the new product in the market. Alternatively, relative product development cycle time is based on a comparison of the elapsed time with planned or expected time to facilitate comparisons across development projects of firms from different industries. Cycle time is also measured as a comparison of the elapsed time of a particular project against schedule, or with the elapsed time of other projects within the firm or those of competitors.

Unit of analysis. A related problem involves the unit of analysis at which accelerated product development is studied: the organization or the project (Kessler and Chakrabarti, 1996). Some variables (e.g., project leader influence, cross-functional teams) that are necessarily measured at the project level may be neither operational nor meaningful at the organizational level. In addition, studies at the organizational level tend to collapse the results of a firm's development projects, obscuring not only the particular characteristics of individual projects, but also the effect of development cycle time on a specific project's success. Analysis at the

project level enables a study to capture unique situational attributes that influence the project processes and outcomes. Not surprisingly, there is a growing consensus in the academic literature that theory and research on accelerated product development should focus at the project level.

Impact. With regard to the conflicting outcomes it has also been argued that the importance of cycle time reduction for new-product success is small compared with other key drivers, and that its effect is thus swamped in the noise. This explanation is certainly a possibility because the financial benefits of accelerated product development, while potentially significant, are likely to be small in comparison with the financial leverage exerted by other factors such as product advantage, the ability of the product to meet customer needs, predevelopment task proficiency, dedicated R&D resources, technological proficiency, and launch proficiency (Henard and Szymanski, 2001). Another line of reasoning is that certain development time thresholds must be exceeded in the concept generation and volume production stages of the development process for any significant effect of cycle time reduction on project success to be detected.

Trade-Offs in Accelerated Product Development

Another explanation for the conflicting outcomes of accelerated product development comes from a theory of trade-offs in new-product development projects (Swink, Talluri, and Pandejpong, 2006). According to this theory, certain practices used to shorten development cycle may be counterproductive in other ways. Most studies on accelerated product development, however, have limited their examination to the extent of development cycle time reduction without any reference to its possible implications on development expense, product performance quality, and/or product profitability. At the same time, these studies suggest that trade-offs may exist between pairs of these development outcomes, requiring that the objectives in accelerated product development should be balanced.

Several studies suggest that the relationship between cycle time and development expenses is U-shaped (Bayus, 1997; Langerak, Hultink, and Griffin, 2008). Reducing development cycle time to below the minimum of the U-shaped curve increases the pressure on financial resources due to higher coordination costs, additional expenses for overtime work, correcting for errors that may have resulted from skipping process steps, and an intensified need for support resources, particularly, among teams involved in accelerated projects. Allowing the new product's cycle time to go above the function's minimum has the similar effect of increasing costs due to decay of know-how, loss of motivation, and the emergence of additional setup costs. An overly loose schedule thus wastes resources because of dissipated efforts and lapses of attention (Langerak, Hultink, and Griffin, 2008).

Researchers have also studied trade-offs among other pairs of development objectives. Calantone and Di Benedetto (2000), for example, provide an analytical model of the link between development cycle time and product performance. They conclude that keeping a new product in development is preferable to the accelerating time to market if the base product performance is low. The trade-off between development cycle time and product performance has also been investigated empirically, and yielded mixed findings (Swink, Talluri, and Pandejpong, 2006). Likewise, support for a development expense–product performance trade-off has been mixed. Finally, timing and performance decisions have been reported to depend on the asymmetries in competitors' market estimates and development efficiencies (Bayus, Jain, and Rao, 1997).

In summary, the literature identifies many mixed findings on the hypothesized trade-offs between development project objectives, suggesting the need for a broader theory that explains the nature of the trade-offs, and empirical research that addresses multiple project outcomes and relevant boundary conditions (e.g., maximum development expenses or minimum product performance) in a more holistic manner.

Acceleration Techniques

Within the literature on accelerated product development, a notable number of studies have

sought to identify those acceleration techniques that new-product teams actually use to reduce development time. The earliest framework for the use of techniques was developed by Millson, Raj, and Wilemon (1992) who suggested a hierarchy of acceleration approaches. Each approach is composed of similar techniques aimed at simplifying development operations, eliminating unnecessary development activities, paralleling development activities, eliminating delays in the process, and speeding up development operations. They concluded that a firm's time to develop new products can be significantly reduced to the extent these five acceleration approaches are employed in a thoughtful manner. In similar vein, Langerak, Peelen, and Nijssen (1999) specified a set of 50 individual acceleration techniques that development teams can adopt to achieve cycle time reduction. They formed nine approaches to shorten development time by clustering similar techniques.

In a follow-up study, Langerak and Hultink (2005) investigated the effect of each cluster on development cycle time and new-product profitability. The results revealed five out of nine acceleration approaches (i.e., supplier involvement, lead user involvement, speeding up activities and tasks, training and rewarding of employees, and simplifying organizational structure) to be associated with reduced development cycle times. Implementing support systems and techniques and stimulating interfunctional coordination, on the other hand, were found to be linked to increased cycle times. The results further showed that three approaches (i.e., lead user involvement, training and rewarding of employees, and an emphasis on the customer) improve new-product profitability, while three other approaches (i.e., speeding up activities and tasks, reduction of parts and components, and implementation of support systems and techniques) decrease new-product profitability. A closer inspection of these results reveals that the use of only two approaches (i.e., lead user involvement and training and rewarding of employees) simultaneously reduces development time and improve new-product profitability. The adoption of one approach (i.e., speeding up activities and tasks) reduces development time to the detriment of profitability, an observation consistent with theories that suggest a trade-off between development time and new-product profitability.

Risks in Cycle Time Reduction

Crawford (1992) argues that the existence of trade-offs also imply risks associated with the practices aimed at cycle time reduction. A firm may focus on incremental new products at the expense of true breakthroughs and necessary information search steps may be sacrificed. Launching a new product too early, particularly when the product is qualitatively not ready for the market, is a notable risk. An overly tight time schedule also raises the probability of mistakes being made because it pushes the development project beyond the limits of R&D, engineering, and marketing capabilities. There are also people costs involved in managing cross-functional teams; time budget cuts can constrain the amount of innovation achieved; and a speeded-up team can consume an inordinate share of firm resources. Furthermore, team members in a functionally organized firm may be under work stress not only because of divided loyalties between their functional area and the development team, but also because of the increase in their workload as a consequence of being a part of the development team. In light of these risks, some researchers have suggested that timely development of a new product, given a new product's window of opportunity, is more important than achieving the shortest development cycle time possible.

Market window of opportunity. The majority of existing work has not taken the new product's window of opportunity into account while investigating the effect of cycle time reduction on the success of the product. The concept of window of opportunity is, however, important because the sales of a new product that is launched ahead or behind its time will suffer (Calantone and Di Benedetto, 2000). Research suggests that by rushing a new product to market, a firm risks facing a strategic window yet to open, because in the early stages of the new product's life cycle there is likely to exist incompatibility with customers' existing way of doing business, and a greater perceived risk of adoption. Moreover, cycle time reduction may negatively affect

product performance. Conversely, by taking too long to develop a new product, a firm may miss the window of opportunity. Customers already exposed to existing products are not likely to postpone their purchase decision, especially if competitors have already introduced similar new products. Contemporary research even suggests that the firm's ability to get the market-entry timing right, in view of the market window, is more important for new-product profitability than reduced cycle time, although it should be kept in mind that market-entry timing is contingent upon the completion of the new product's development cycle (Langerak, Hultink, and Griffin, 2008).

Conclusion

This discussion of accelerated product development provides scholars with a theoretical foundation for rigorous, empirical research. The challenge is, of course, to build on prior studies to investigate and further clarify some of the issues without losing sight of the complex relationships pertaining to the context, antecedents, and outcomes of development cycle time reduction. In doing so, researchers should consider the need for consistency in the unit of analysis and the measurement of development cycle time. From a practitioner's perspective, this discussion is useful in its clarification of the merits of cycle time reduction and the situations in which it is most appropriate, delineation of the ways in which interventions can be applied, explanation of the risks, trade-offs and pitfalls of development time reduction, and exploration of the competitive implications of accelerated product development.

Acknowledgment

The author thanks Pinar Cankurtaran for her insightful and helpful comments.

Bibliography

Ali, A. (2000) The impact of innovativeness and development time on new product performance for small firms. *Marketing Letters*, **11** (2), 151–163.

Bayus, B.L. (1997) Speed-to-market and new product performance trade-offs. *Journal of Product Innovation Management*, **14** (6), 485–497.

Bayus, B.L., Jain, S., and Rao, A.G. (1997) Too little, too early: introduction timing and new product performance in the personal digital assistant industry. *Journal of Marketing Research*, **34**, 50–63.

Calantone, R.J. and Di Benedetto, C.A. (2000) Performance and time to market: accelerating cycle time with overlapping stages. *IEEE Transactions on Engineering Management*, **47** (2), 232–244.

Chen, J., Reilly, R.R., and Lynn, G.S. (2005) The impacts of speed-to-market on new product success: the moderating effects of uncertainty. *IEEE Transactions on Engineering Management*, **52** (2), 199–212.

Crawford, C.M. (1992) The hidden costs of accelerated product development. *Journal of Product Innovation Management*, **9** (3), 188–199.

Gerwin, D. and Barrowman, N.J. (2002) An evaluation of research on integrated product management. *Management Science*, **48** (7), 938–953.

Griffin, A. (1997) Modeling and measuring product development cycle time across industries. *Journal of Engineering and Technology Management*, **14** (1), 1–24.

Griffin, A. (2002) Product development cycle time for business-to-business products. *Industrial Marketing Management*, **31** (4), 291–304.

Henard, D.H. and Szymanski, D.M. (2001) Why some new products are more successful than others. *Journal of Marketing Research*, **38** (3), 362–375.

Kessler, E.H. and Chakrabarti, A.K. (1996) Innovation speed: a conceptual model of context, antecedents, and outcomes. *Academy of Management Review*, **21** (4), 1143–1191.

Langerak, F. and Hultink, E.J. (2005) The impact of new product development acceleration approaches on speed and profitability: lessons for pioneers and fast followers. *IEEE Transactions on Engineering Management*, **52** (1), 30–42.

Langerak, F., Hultink, E.J., and Griffin, A. (2008) Exploring mediating and moderating influences on the links among cycle time, proficiency in entry timing and new product profitability. *Journal of Product Innovation Management*, **25** (4), 370–385.

Langerak, F., Peelen, E., and Nijssen, E.J. (1999) A laddering approach to the use of methods and techniques to reduce the cycle time of new-to-the-firm products. *Journal of Product Innovation Management*, **16** (2), 173–182.

Millson, M.R., Raj, S.P., and Wilemon, D. (1992) A survey of major approaches for accelerating new product development. *Journal of Product Innovation Management*, **9** (1), 53–69.

Swink, M., Talluri, S., and Pandejpong, T. (2006) Faster, better, cheaper: a study of NPD project efficiency and performance tradeoffs. *Journal of Operations Management*, **24** (5), 542–562.

product life cycle

Peter N. Golder and Gerard J. Tellis

Introduction

The product life cycle (PLC) is a central, enduring framework in marketing and many other fields including economics, management, production, logistics, engineering, political science, and travel and tourism. In this article, we present its theoretical underpinnings, metrics for its key turning points, and general findings of its characteristics.

The PLC describes the sales history of a new product in terms of four stages: introduction, growth, maturity, and decline. The concept was developed initially in the early 1950s, triggering extensive empirical research in the 1960s (Rink and Swan, 1979). Most of this research sought to validate the PLC in one or more groups of products. Yet, this research did not yield generalizations about the PLC, and the primary researchers differed in their conclusions about the usefulness of the PLC (Polli and Cook, 1969; Buzzell, 1969). Three primary factors underlie these different conclusions:

- *Level of aggregation*: The PLC can apply to various levels of product aggregation: the product class, form, or brand. Different studies tested the PLC at different levels. This made comparisons across studies difficult, especially if different patterns hold at different levels of aggregation. The PLC concept may be especially inappropriate for brand-level analysis. We agree with Lambkin and Day (1989) that the product category level is most appropriate for the PLC concept.
- *Research focus*: Past studies focused on products or brands in substantially different groups of products, such as food, personal care, consumer durables, and industrial products. The PLC is more likely to have distinct patterns *within* such groups of products, particularly, new consumer durables, which are significant improvements over existing products.
- *Nature of the concept*: As with other concepts in the social sciences, the PLC is likely to be probabilistic rather than deterministic. Past empirical studies tended to highlight unique differences in life cycles as deviations from a deterministic pattern, rather than emphasizing probabilistic commonalties.

Definitions

We define a *product category* as a group of products without close substitutes that, from the consumer's viewpoint, fulfill a distinct need. For example, refrigerators, CD players, and camcorders are product categories. Each of these categories may encompass multiple brands that are close substitutes of each other. A product form is a subcategory of products that satisfy particular consumer needs within the product category. Product form is a level of aggregation between product category and brand. For example, VHS, VHS-C, 8 mm, and mini-DVD are different forms within the product category of camcorders. Product forms typically meet the needs of a distinct segment of consumers in the market for camcorders.

We define the four standard stages of the PLC, as follows: (i) *introduction* is the period from a new product's commercialization until its takeoff; (ii) *growth* is the period from a new product's takeoff until its slowdown in sales; (iii) *maturity* is the period from a product's slowdown until sales begin a steady decline because of the emergence of a new product that satisfies essentially the same need; and (iv) *decline* is the period of steadily decreasing sales until a product's demise due to complete substitution by a superior new product.

In addition, we define three key events that mark the beginning and end of the first two stages: (i) *commercialization* is the point at which a new product category is first sold to consumers; (ii) *takeoff* is the point of transition from the introduction to the growth stage of the PLC. It is the first dramatic and sustained increase in product category sales (see TAKEOFF). (iii) *Slowdown* is the point of

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

transition from the growth stage to the maturity stage of the PLC. The slowdown signals the beginning of a period of level, slowly increasing, or temporarily decreasing product category sales.

Measurement of commercialization is fairly straightforward, although there are often different starting years for businesses and individual customers (Golder and Tellis, 1997; Golder, Shacham, and Mitra, 2009). A common metric for measuring takeoff is the first year in which a product's growth rate relative to its previous year's unit sales is higher than a predetermined threshold (Golder and Tellis, 1997). Other studies measure growth relative to market penetration, which is more appropriate when investigating commonalities across countries (Chandrasekaran and Tellis, 2008; Tellis, Stremersch, and Yin, 2003). A measure of slowdown is the first year of two consecutive years after takeoff, in which sales are lower than the highest previous sales. The measures of takeoff and slowdown match a visual inspection of the sales curves in nearly all product categories.

An Integrated Theory of the Product Life Cycle

Our integrated theory draws on research in marketing, diffusion, behavioral economics, and technology management. The PLC is primarily the confluence of two general underlying forces: consumer demand and producer supply. These two forces work simultaneously and interactively to give the life cycle its shape.

Consumer demand. Consumer demand is the primary force driving the PLC. In particular, three important components of demand play a role in shaping the PLC: diffusion, affordability, and informational cascades.

Diffusion. Diffusion explains how the knowledge of a new product evolves through a population (*see also* DIFFUSION OF INNOVATION). The idea has roots in the work of Everett Rogers, and has been extensively documented in marketing (Lambkin and Day, 1989; Mahajan, Muller, and Bass, 1990; Rink and Swan, 1979; Rogers, 2003). This stream of research suggests that knowledge of innovations and their benefits diffuse in a population, from segment to segment, and from user to user within segments, by media reports, word of mouth, observation, and imitation. The reason it diffuses this way is because these segments differ in their propensity to innovate. In particular, certain factors, such as age, education, mobility, and social desirability, influence this propensity to adopt innovations. As a result of these differences, various consumer segments are likely to adopt innovations sequentially. A great body of research reports that, typically, the segments that adopt the innovation very early or very late are small, while those that adopt it in between these groups are large. As a result, a plot of adoptions against time gives the typical normal or bell-shaped PLC curve. By adding repeat sales to sales from adoptions, one gets the total sales curve. If adoptions typically follow a normal distribution, then the early part of the life cycle of a new product will have an S-shape. The length and shape of the maturity stage of the PLC depends increasingly on replacement purchases. For fad and fashion products, this stage may not last long. For durables, this stage can be very long.

Affordability. The actual purchase of a product occurs when consumers find that the benefits of the product are equal to or higher than the price of the product. Thus, affordability plays an important role in takeoff and growth. Typically, information about the new product diffuses much more rapidly through a population than does the purchase of the product itself. Therefore, far more people want new products than are willing to buy them at their current prices. Because of technological improvements and manufacturing efficiencies, new products improve rapidly in benefits to consumers, while their prices come down. The prices of new products decrease substantially during the introduction stage (Golder and Tellis, 1997). As prices continue to decline during the growth stage, affordability is enhanced and sales push higher.

Informational cascades. Informational cascades describe how people converge on adopting a new product (or behavior) with

increasing momentum and declining individual evaluation of the merits of the new product, because of their tendency to derive information from the behavior of prior adopters (Bikhchandani, Hirshleifer, and Welch, 1992). The essence of informational cascades is that even though individuals make decisions on the basis of their own private information, their decisions are also influenced by other people's decisions. As people adopt a new product on the basis of its merits, their adoption provides a signal to nonadopters. Some of these nonadopters go on to adopt the new product, at least partially influenced by the behavior of the previous adopters. As the number of adopters increases, they provide an increasingly strong signal to nonadopters, who then adopt in increasing numbers. Once information derived from the decisions of others begins to outweigh an individual's private valuation, the process begins to increase in momentum or cascade toward conformance in the behavior of more and more buyers.

At this point, new adopters reveal no additional private information to the market. Thus, the mass conformance or cascade to a particular behavior is based on the initial decisions of a small number of adopters, rather than the cumulative decisions of all adopters. This feature makes a cascade quite fragile. It is easily triggered or reversed by new information that affects the decisions of a small number of people. Therefore, cascades can depress pre-takeoff sales, sharpen the takeoff of new products, exaggerate product growth, and reverse sales growth when maturity first sets in. We elaborate on these points. Since most individuals do not purchase a new product during the introduction stage, most consumers use this information to decide likewise. This decision may occur even when, for some individuals, private information alone would have led to purchasing the new product. As a result, early sales of a new product could be depressed or time to takeoff could be longer than seems justified by the potential utility of the new product (Golder and Tellis, 1997; Tellis, Stremersch, and Yin, 2003).

On the other hand, once a sizable segment of consumers adopts the new product, other consumers use this information to decide to purchase the new product as well. These purchases occur even when their own private valuation could have meant they should not purchase the new product. In such a positive cascade, more consumers buy than might be expected on a strict cost–benefit analysis. This cascade of consumers to adopt a new product is likely to end somewhere. Potential triggers include a decline in the growth rate of the new product's benefits, the announcement of a rival technology, or a change in the economic environment, such as an increase in interest rates or the onset of a recession. Any of these may be small shocks to the system. However, in the presence of a cascade in favor of the new product, any of them might be enough to reverse the sales trend. Some informed consumers might decide to wait before purchasing the new product. Other consumers may have already bought the product before their private valuations justified such a purchase. Once some consumers delay their purchases and other consumers become aware of these decisions, a negative cascade begins. As a result, the slowdown in market sales might not be a gradual flattening of the sales curve, but a drop in sales at the onset of maturity.

Producer supply. Producer supply is the second major force driving the PLC. In particular, two important components of supply play a role in shaping the PLC: technology evolution and market competition.

Technology evolution. While demand-based forces are widely recognized in marketing, supply-based forces, such as technology, are considered less frequently. However, technology-based life cycles have a strong theoretical foundation and are particularly relevant to consumer durables. Therefore, some progress on PLCs may be possible by considering these supply-side factors (Lambkin and Day, 1989). The role of technology in the PLC can be succinctly described by the phenomenon of the S-curves of technological change (Chandy and Tellis, 2000; Foster, 1986). Product improvements are commonly driven by technological evolution that proceeds along a series of successive S-curves (*see* TECHNOLOGY S-CURVE). The S-curve emerges because an

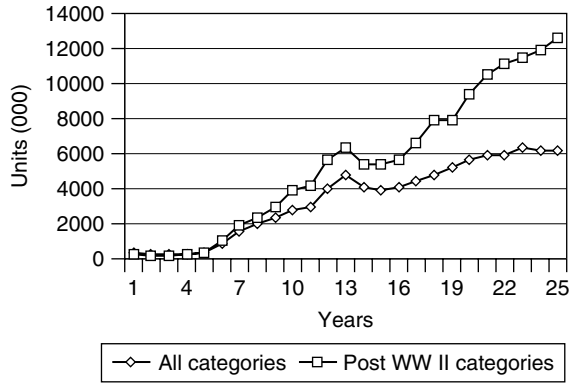


Figure 1 Average sales history.

existing technology improves slowly in benefits to consumers when first introduced, then rapidly as research on the technology advances, and finally slowly again, as research reaches the limits of that technology. The pattern of S-curves plays a critical role in the PLC. However, regularity in S-curves is itself questionable while the life cycle of the product does not mimic the S-curve of its underlying technology (Sood and Tellis, 2005). For example, the takeoff of the new product occurs when the new technology surpasses the old technology and not at the takeoff of the new technology.

Market competition. The final force driving the PLC has been studied during the introduction and early growth stages. One of the main determinants of sales takeoff is found to be the number of competitors in the market. Interestingly, there is a takeoff in the number of competitors that tends to precede the takeoff in sales. More competitors are likely to drive improvements in quality (both actual and perceived) and make a contribution to sales gains that may supersede even the positive impact of declining prices (Agarwal and Bayus, 2002).

Empirical Regularities in the PLC of Consumer Durables

We now report observed regularities in the PLC based on a sample of 30 new consumer durables consisting of household appliances and consumer electronics introduced between 1900 and 1990 (Golder and Tellis, 2004).

We use sales data in each category to construct a "typical" or average PLC for all categories and for the subset of categories introduced after World War II (see Figure 1). The three key events enable us to demarcate the stages of new-product sales and present descriptive results on growth rates and unit sales. Our analysis of sales after the slowdown leads us to divide the maturity stage into two substages, early maturity and late maturity. Early maturity begins with the year of the sales slowdown and continues until sales grow to the previous local peak. Late maturity begins with the first year that sales are higher than the local peak and continues until a product's sales begin to fall steadily during the decline stage. Only a few of our categories have reached the decline stage, so we do not report findings for this stage.

Figure 1 confirms the importance of properly identifying takeoff and slowdown because the average life cycle changes so distinctively at these points. In particular, after slowdown, sales take several years to surpass the previous local peak (Goldenberg, Libai, and Muller, 2002).

Characterization of stages and events. Table 1 contains the growth rates during the stages and at the key events of the PLC. Growth rates vary dramatically over the PLC. While growth is high during the introduction, the base is relatively small; so these percentages do not represent substantial increases in unit

Table 1 Growth rate (%).

	Introduction[a]	*Takeoff*	*Growth*	*Slowdown*	*Early Maturity*[b]	*Late Maturity*
Total Sample	**31**	**428**	**45**	**−15**	**−25**	**3.7**
Standard deviation	39	995	39	13	15	5.7
Median	15	207	32	−12	−22	2.9
10–90th percentile	6–74	33–491	15–92	–33–(–4)	–42–(–9)	–2.2–10.9
Sample size	9	30	30	25	25	19
Pre WW II	**19**	**181**	**39**	**−14**	**−29**	**3.2**
Standard deviation	24	154	44	11	10	5.8
Median	12	182	23	−14	−28	2.5
10–90th percentile	5–42	31–429	14–79	–25–(–5)	–42–(–17)	–2.7–8.5
Sample size	5	14	14	14	14	14
Post WW II	**45**	**645**	**50**	**−15**	**−21**	**5.0**
Standard deviation	52	1337	34	15	19	5.9
Median	22	249	44	−9	−14	4.4
10–90th percentile	17–93	84–907	18–81	–37–(–4)	–36–(–5)	–0.1–11.1
Sample size	4	16	16	11	11	5

[a] Nine categories with sales data for three years before takeoff and base sales of at least 15 000 units.
[b] Minimum sales after slowdown relative to sales at local peak.

sales. However, at takeoff, the market for these new products changes dramatically with an average sales increase of over 400%. During the growth stage, sales increase at 45% per year, on average. In addition, the higher base of sales after takeoff means that the increases in unit sales are quite substantial (*see also* GROWTH STRATEGIES).

At slowdown, the sustained period of rapid growth suddenly reverses into a 15% decline. Our theoretical discussion of informational cascades helps to shed some insight into the sales reversal at slowdown. After slowdown, sales tend to decline for a few years, reaching a minimum of 25% below the local sales peak prior to slowdown. During late maturity, sales surpass the previous local peak in sales but tend to grow at a much slower rate (3.7%). Sales growth during maturity seems to reflect increases in the total economy (*see also* MANAGING MATURE PRODUCTS).

Table 2 contains the mean unit sales at three key events: commercialization, takeoff, and slowdown. These averages provide benchmarks for judging the success of today's new products (*see also* NEW-PRODUCT FORECASTING).

Table 3 presents results on market penetration at takeoff and slowdown. Penetration at takeoff varies significantly between older and newer categories, but average penetration at slowdown does not vary.

Information about relative prices is presented in Table 4. Larger price reductions in more recent periods may be because of stronger experience effects associated with larger national and international markets, more electronic components, and managers exploiting these experience effects more often.

Results on the stage durations are presented in Table 5. The duration of the growth stage has not been shortening over time. However, the durations of introduction and early maturity

Table 2 Unit sales (thousands).

	Commercialization	*Takeoff*	*Slowdown*
Total Sample	**34**	**902**	**5839**
Standard deviation	41	754	6904
Median	5	654	3828
10—90th percentile	2.9–83	317–1690	1296–11 496
Sample size	9	30	25
Pre WW II	**3.0**	**728**	**3243**
Standard deviation	—	414	1700
Median	—	574	3560
10—90th percentile	—	349–1336	1266–5300
Sample size	1	14	14
Post WW II	**38**	**1053**	**9142**
Standard deviation	42	948	9444
Median	25	802	5000
10—90th percentile	2.9–88	313–2464	2585–21 697
Sample size	8	16	11

Table 3 Market penetration (%).

	Takeoff	*Slowdown*
Total Sample	**2.9**	**34.2**
Standard deviation	3.4	20.3
Median	1.4	36.6
10–90th percentile	0.4–6.8	12.0–51.0
Sample size	30	25
Pre WW II	**4.3**	**32.5**
Standard deviation	4.3	14.3
Median	2.2	36.6
10–90th percentile	1.0–9.7	12.5–47.3
Sample size	14	14
Post WW II	**1.7**	**36.4**
Standard deviation	1.7	26.8
Median	1.1	38.2
10–90th percentile	0.4–4.2	12.0–55.0
Sample size	16	11

Table 4 Price relative to commercialization price.

	Takeoff	*Slowdown*
Total Sample	**0.71**	**0.44**
Standard deviation	0.20	0.26
Median	0.76	0.44
10–90th percentile	0.42–0.97	0.15–0.78
Sample size	30	25
Pre WW II	**0.80**	**0.56**
Standard deviation	0.18	0.25
Median	0.83	0.54
10–90th percentile	0.56–0.98	0.26–0.84
Sample size	14	14
Post WW II	**0.63**	**0.30**
Standard deviation	0.20	0.19
Median	0.64	0.31
10–90th percentile	0.41–0.85	0.05–0.50
Sample size	16	11

have been decreasing over time. We recommend that interested readers consult other studies on the duration of PLC stages (Bayus, 1994, 1998; Qualls, Olshavsky, and Michaels, 1981).

Conclusion

Our theoretical justification for the PLC explains why it is a stable, recurring marketing phenomenon. The commonality in descriptive statistics for 30 new consumer durables lends additional support to this view. Moreover, these statistics provide benchmarks for managers and researchers of today's new products. On the basis of our theory and results, managers would need to anticipate an extended pre-takeoff period, a sharp takeoff, and then a slowdown in sales after about eight years of rapid growth. By

Table 5 Duration of stages (years).

	Introduction	*Growth*	*Early Maturity*
Total Sample	**10.4**	**8.4**	**5.5**
Standard deviation	11.4	4.4	3.7
Median	8	7	5
10–90th percentile	1.9–20	3.0–15	2.0–8.8
Sample size	30	25	19
Pre WW II	**16.9**	**8.6**	**6.4**
Standard deviation	13.9	4.5	3.8
Median	14	7.5	5
10–90th percentile	5.2–24.0	3.3–14.1	3.2–11.2
Sample size	14	14	13
Post WW II	**4.8**	**8.1**	**3.5**
Standard deviation	3.4	4.4	2.7
Median	4	7	2
10–90th percentile	1.0–9.5	3.0–14.0	1.5–7.0
Sample size	16	11	6

anticipating and planning for these events, managers can make more timely investments in manufacturing, sales force, inventory, and marketing.

Bibliography

Agarwal, R. and Bayus, B.L. (2002) The market evolution and sales takeoff of product innovations. *Management Science*, **48**, 1024–1041.

Bayus, B.L. (1994) Are product life cycles really getting shorter? *Journal of Product Innovation Management*, **11**, 300–308.

Bayus, B.L. (1998) An analysis of product lifetimes in a technologically dynamic industry. *Management Science*, **44**, 763–775.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992) A theory of fads, fashion, custom, and cultural change as information cascades. *Journal of Political Economy*, **100** (5), 992–1026.

Buzzell, R.D. (1969) Product Life Cycles, Report No. 69-108, Marketing Science Institute, sections one and four.

Chandrasekaran, D. and Tellis, G.J. (2008) The global takeoff of new products: culture, wealth, or vanishing differences. *Marketing Science*, **27** (5), 844–860.

Chandy, R. and Tellis, G.J. (2000) The incumbent's curse? Incumbency, size and radical product innovation. *Journal of Marketing*, **64** (3), 1–17.

Day, G.S. (1981) The product life cycle: analysis and application issues. *Journal of Marketing*, **45**, 60–67.

Foster, R. (1986) *Innovation: The Attacker's Advantage*, Summit Books, New York.

Goldenberg, J., Libai, B., and Muller, E. (2002) Riding the saddle: how cross-market communications can create a major slump in sales. *Journal of Marketing*, **66** (2), 1–16.

Golder, P.N., Shacham, R., and Mitra, D. (2009) Innovations' origins: when, by whom, and how are radical innovations developed? *Marketing Science*, **28** (1), 166–179.

Golder, P.N. and Tellis, G.J. (1997) Will it ever fly? Modeling the takeoff of really new consumer durables. *Marketing Science*, **16** (3), 256–270.

Golder, P.N. and Tellis, G.J. (1998) Beyond diffusion: an affordability model of the growth of new consumer durables. *Journal of Forecasting*, **17**, 259–280.

Golder, P.N. and Tellis, G.J. (2004) Growing, growing, gone: cascades, diffusion, and turning points in the product life cycle. *Marketing Science*, **2**, 207–218.

Lambkin, M. and Day, G.S. (1989) Evolutionary processes in competitive markets: beyond the product life cycle. *Journal of Marketing*, **53**, 4–20.

Mahajan, V., Muller, E., and Bass, F.M. (1990) New product diffusion models in marketing: a review and directions for research. *Journal of Marketing*, **54**, 1–26.

Polli, R.E. and Cook, V. (1969) Validity of the product life cycle. *Journal of Business*, **42**, 385–400.

Qualls, W., Olshavsky, R.W., and Michaels, R.E. (1981) Shortening of the PLC – an empirical test. *Journal of Marketing*, **45** (Fall), 76–80.

Rink, D.R. and Swan, J.E. (1979) Product life cycle research: a literature review. *Journal of Business Research*, **78**, 219–242.

Rogers, E.M. (2003) *Diffusion of Innovations*, 5th edn, The Free Press, New York.

Sood, A. and Tellis, G.J. (2005) Technological evolution and radical innovations. *Journal of Marketing*, **69** (3), 152–168.

Tellis, G.J., Stremersch, S., and Yin, E. (2003) The international takeoff of new products: economics, culture, and country innovativeness. *Marketing Science*, **22** (Spring), 188–208.

portfolio management

Scott J. Edgett

Introduction

Product innovation is vital to the modern corporation. Driven by rapidly advancing technologies, globalization of markets, and increasing competition at home and abroad, effective product innovation is critical to successful corporate performance.

A vital question in product innovation is: How should corporations effectively allocate their research and development (R&D) and new-product resources? Portfolio management is all about maximizing resources to achieve strategic product innovation objectives. Executives who optimize their R&D investments by defining the right product innovation strategy for the firm, by selecting winning new-product projects, and by achieving the ideal balance of projects, will come out ahead in the long run. These executives practice effective portfolio management for product innovation.

Portfolio Management Defined

The portfolio decision-making process encompasses, or overlaps, a number of decision-making processes within the business. Portfolio decisions include periodic and holistic reviews of the total portfolio of all projects, making Go/Kill decisions on an ongoing basis, and implementing the product innovation strategy for the business. This includes strategic resource allocation decisions within selected arenas.

Portfolio management is operationalized on both the strategic and tactical levels. The portfolio decision process is characterized by uncertain and changing information, dynamic opportunities, multiple goals and strategic considerations, interdependence among projects, and multiple decision makers (Cooper, Edgett, and Kleinschmidt, 1997a, 2002a, 2002b). Portfolio prioritization is a dynamic decision process. A business's list of active new-product (and development) projects is constantly updated and revised. In this process, new projects are evaluated, selected, and prioritized. Existing projects may be accelerated, killed, or de-prioritized, with resources reallocated to active projects as required (Figure 1).

Challenges to Good Portfolio Management Practices

There are four unique facets to good portfolio management that make it one of the more challenging decision-making functions in an organization:

- Portfolio management for product innovation deals with future events and opportunities; thus, much of the information required to make project selection decisions is at best uncertain, and at worst unreliable.
- The decision environment is dynamic: The status and prospects for projects in the portfolio are ever changing, as markets shift and new information becomes available.
- Projects in the portfolio are at different stages of completion, yet all projects compete against each other for resources. Therefore, comparisons must be made between projects based on information that differs in quality and quantity.
- Finally, resources available to be allocated across projects are limited: A decision to fund one project may mean that resources must be taken away from another; and resource transfers between projects are usually not totally seamless.

Three Main Goals in Portfolio Management

How should one go about setting up a portfolio management system? There are three main goals that an effective portfolio management system should achieve:

- *Strategic alignment*. The main goal is to ensure that, regardless of all other considerations, the final portfolio of projects truly reflects the company's innovation strategy: that all projects are "on strategy," support the strategy, and/or are critical components of the strategy; and, that the breakdown of spending across projects, areas, and markets, is directly tied to the business strategy.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

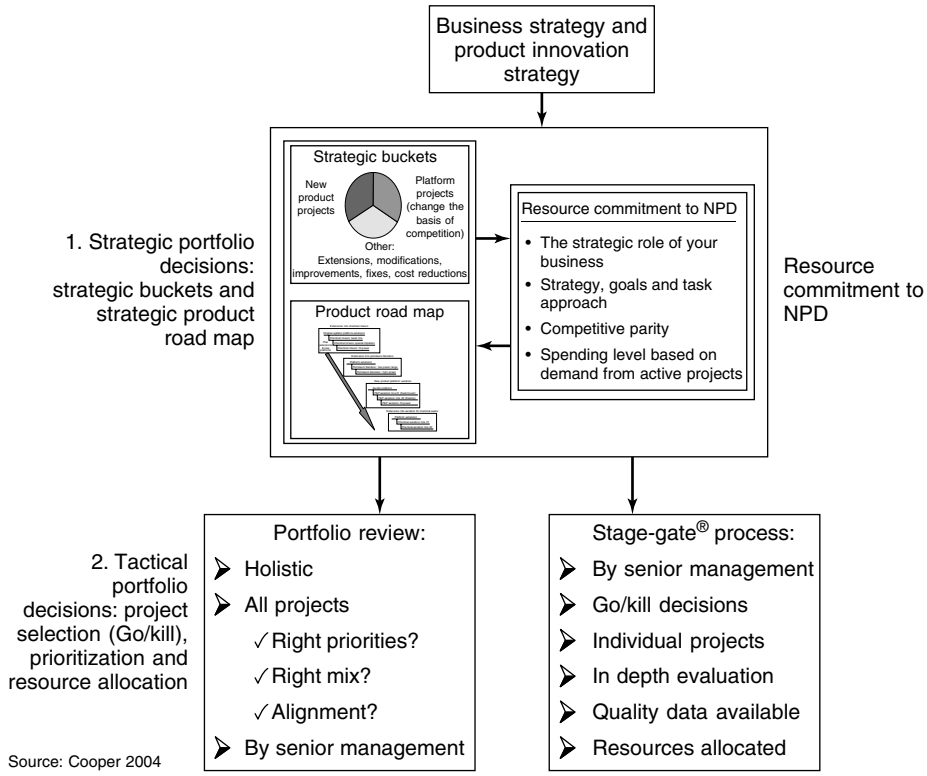


Figure 1 The portfolio management system and its elements (reproduced from Cooper, R.G. (2004) Product Leadership: Pathways to Profitable Innovation, 2nd Edition. Cambridge, MA: Basic Books. © The Perseus Book Group).

- *Maximize the value of the portfolio*. Here the goal is to allocate resources to maximize value for a given spending level. In other words, projects are selected to maximize the sum of the commercial value of all active projects in the pipeline. This can be measured by a business objective such as net present value (NPV), return on investment (ROI), or likelihood of success.
- *The right balance of projects*. The goal is to achieve the desired balance of projects in terms of a number of parameters to ensure strategic alignment. For example, a company might seek the right balance in terms of long-term projects versus short-term ones; or high-risk versus lower-risk; or across various markets, technologies, product categories, and project types (e.g., new products, improvements, cost reductions, maintenance and fixes, and fundamental research) (*see also* INNOVATION TYPOLOGIES).

To achieve these three key goals, portfolio management is separated into two distinct but related levels: strategic and tactical. This permits different decisions to be addressed at their appropriate degree of intensity within the organization (e.g., whether the issue is of a strategic nature or of a resource allocation nature).

Strategic Portfolio Management

Strategic portfolio decisions answer the following questions:

- Directionally, where should the business deploy its product innovation resources (people and funds)?

- How should resources be split across project types, markets, technologies, or product categories?
- On what major initiatives or new platforms should the business concentrate its resources?

Typical aids to answer these questions are (i) creating a strategic product roadmap; (ii) defining strategic buckets (or arenas); and (iii) determining the total resource commitment necessary to achieve innovation goals (money and people).

Creating the strategic product road map. The development of a product road map flows logically from the product innovation strategy. Delineating the major initiatives required for the road map includes the following:

- *Strategic assessment*. Sometimes the mere specification of a strategic arena as a top priority leads logically to a list of products and projects that are necessary to be successful in that arena. For example, a major food-products company identified "healthy snack" as a priority strategic arena (the company already sold a few products in this sector, but was a minor player). As soon as this happened, the products required to be successful in this market segment became self-evident. As a result, the development programs needed to generate these products fell into a logical sequence in a product road map.
- *Portfolio review of existing products*. Here, a hard look at the business's current product offerings is undertaken, and decisions are made as to which products should be cut, continued, or replaced. Forecasts of products' life cycles often reveal the need and timing for replacement products and, even new platforms. Additionally, gaps in the product lines are identified (*see* PRODUCT-LINE STRATEGIES). Then, place marks are inserted in the product road map for these development actions. This exercise is undertaken periodically to keep the product line fresh and complete.
- *Competitive analysis*. Where are the business's current and future products and product lines relative to competitors (*see* COMPETITIVE ADVANTAGE): this is evaluated by assessing competitors' current and probable future offerings, where they have advantages and where the gaps are. This exercise often points to the need for new products.
- *Technology trends assessment*. Here one forecasts the types and timing of new technologies and the new platform developments that will be required as a result (*see* PRODUCT PLATFORMS). For example, the advent of each new cell-phone technology signals a host of development projects within cell-phone manufacturing firms, as well as within service providers.
- *Market trends assessment*. This is also a forecasting exercise that looks at major market trends and shifts (*see also* NEW-PRODUCT FORECASTING). In this exercise, one is often able to pinpoint specific initiatives that must be undertaken. An example is shifting purchasing patterns in favor of more environmentally friendly products.

Strategic buckets. When translating the business's strategy into strategic portfolio decisions, a major challenge is determining spending breakdown or deployment; that is, where does senior management need to spend its resources when it comes to product innovation – on what types of projects, and in what product, market or technology areas? And how much do they need to spend in each area? The strategic-buckets model operates from the simple principle that implementing strategy equates to spending money on specific projects. Thus, operationalizing strategy really means "setting spending targets." Strategic buckets can be used with, or instead of, the product road map.

This method begins with the business's strategy, and requires senior management to make choices that enable the creation of "envelopes of resources" or "buckets." Existing projects are categorized into buckets; then, actual spending is reviewed for consistency with desired spending for each bucket. Finally, projects are prioritized within buckets to arrive at the ultimate portfolio of projects – one that mirrors management's strategy for the company.

Some common ways to categorize buckets are

- *Along strategic goals*. Management splits resources across the specified strategic goals. For example, what percentage will be spent on defending the base? On diversifying? On extending the base?
- *Across arenas*. This means defining the spending splits across strategic arenas defined in the business strategy. Arenas are generally product, market, or technology areas where the business wishes to focus its new-product efforts.
- *Along product lines*. Resources are divided among product lines. A plot of product-line locations on the product life-cycle curve is used to help determine this split.
- *Along types of projects*. Buckets may be allocated according to the types of projects.
- *By technologies or technology platforms*. Spending splits can be made across technology types (e.g., base, key, pacing, and embryonic technologies), or across specific technology platforms.
- *By familiarity matrix*. What should be the split of resources to different types of markets and to different technology types in terms of their familiarity to the business? Some companies use the popular "familiarity matrix" – technology newness versus market newness – to help split resources.
- *By geography*. What proportion of resources should be spent on projects aimed largely at North America? At Latin America? At Europe? At Asia-Pacific?, Or globally?
- *By stage of development*. Some businesses distinguish between early stage projects and projects in development and beyond. Two buckets are created, one for development projects, the other for early stage projects.

The major question with strategic buckets is, "How does a company obtain the optimal split in project types?" Breaking down new products and projects by type helps predict a business's product innovation performance. For example, too much emphasis on short term, small projects might point to an underachieving business (see Cooper, Edgett, and Kleinschmidt (2005) for additional examples and data). So, ideally a business's product innovation strategy should be reflected in the types of product development it undertakes and in where it invests its product innovation resources.

Determining resource commitment. Once the product road maps and strategic buckets have been determined, the third challenge is to determine what the total resource commitment should be to achieve the innovation goals and to properly allocate funds to the various road maps and strategic buckets. The issue here is that there is seldom enough money to fund all of the potential opportunities. So the first decision is how much resources, in aggregate, are needed to properly fund the R&D initiatives to ensure a reasonable probability of success in achieving the strategic goals. The second set of decisions is to properly allocate the funds among the different road maps and arenas. The process for this will usually require several rounds of allocation that result in a series of prioritization choices being made.

Tactical Portfolio Management

Tactical portfolio decisions focus on individual project selection, but obviously follow from the strategic decisions. Questions typically addressed are as follows: (i) what specific new-product and development projects should the business undertake? (ii) What priority should these individual projects have? (iii) How should the resources be allocated to each project?

Project selection. Two processes that enable project selection are necessary to make effective tactical decisions. Both work in harmony. They are gates and portfolio reviews.

- *Gates*. Embedded within the idea-to-launch new-product framework are Go/Kill decision points called "*gates*" (the most commonly used process worldwide is the Stage-Gate® process – *see* THE STAGE-GATE IDEA TO LAUNCH SYSTEM). Gates allow for in-depth review of individual projects. They provide for Go/Kill, prioritization, and resource allocation decisions. Gates must be part of the portfolio management system.

- *Periodic portfolio review*. Senior management should meet two to four times per year to review the portfolio of all projects. Key issues and questions addressed at these meetings are
  - Are all projects strategically aligned to fit the business' strategy?
  - Does management have the right priorities among projects?
  - Are there some projects on the active list that should be killed or, perhaps, accelerated?
  - Is there the right balance and the right mix of projects?
  - Are there enough resources to undertake all of the projects?
  - Is there sufficiency – if one undertakes these projects, will the business achieve its stated business goals?

Both decision processes – gating and portfolio reviews – are necessary. Note that the gates are project specific and provide a thorough review of each project, in depth and in real time. By contrast, portfolio reviews are holistic: They look at all projects together, but in much less detail. Ideally, if the gates are working, few major decisions or corrective actions are required at the portfolio review. For example, some companies indicate that they only consider projects in aggregate at the portfolio review. However, if the gating process is not working well, then most decisions are made at these quarterly or semiannual portfolio reviews.

Tools for Effective Gate and Portfolio Reviews

Within the gate and portfolio reviews, a number of tools can be used to help achieve portfolio goals – maximize the portfolio's value, achieve the right balance and mix of projects, and ensure strategic alignment – yet not overload the development pipeline.

Maximizing the value of the portfolio. Methods used to maximize value range from financial tools to balanced scorecard models. Each has its strengths and weaknesses (Cooper, Edgett, and Kleinschmidt, 2002a, 2005). The end result of each method is a rank-ordered, or prioritized, list of "go" and "hold" projects, with the projects at the top of the list scoring highest in terms of achieving the desired objectives. The portfolio's value in terms of that objective is thus maximized. There are many methods available to determine the optimal list of projects. Three of the most common approaches are

- Rank order projects using a financial tool such as NPV.
- Rank order projects using a productivity index based on the NPV. (Note: a productivity index is the ratio of what one is trying to maximize divided by the constraining resource.)
- A balanced scorecard or scoring model approach is based on the premise that a more balanced approach to project selection is desired. A variety of criteria may be used to rate the project. Typical criteria include strategic alignment, product and competitive advantage, market attractiveness, ability to leverage core competencies, technical feasibility, and reward versus risk.

Seeking the right balance of projects. A major portfolio goal is balance (a balanced set of development projects in terms of a number of key parameters). The analogy is that of an investment fund, where the fund manager seeks balance in terms of high-risk versus low-risk stocks; and, balance across industries and geographic areas, to arrive at an optimum investment portfolio. Charts are the most common and effective way to display balance in a product innovation portfolio. These visual representations include portfolio maps or bubble diagrams as well as more traditional pie charts and histograms.

The graphical output can illustrate a number of themes depending on key attributes on which management is trying to achieve balance; for example, risk versus reward, newness to the firm, resource breakdown by project type, markets, products and/or technologies, market timing, and cash flow.

Implementing A Systematic Portfolio Management Process

Top-performing businesses in product development have implemented systematic portfolio

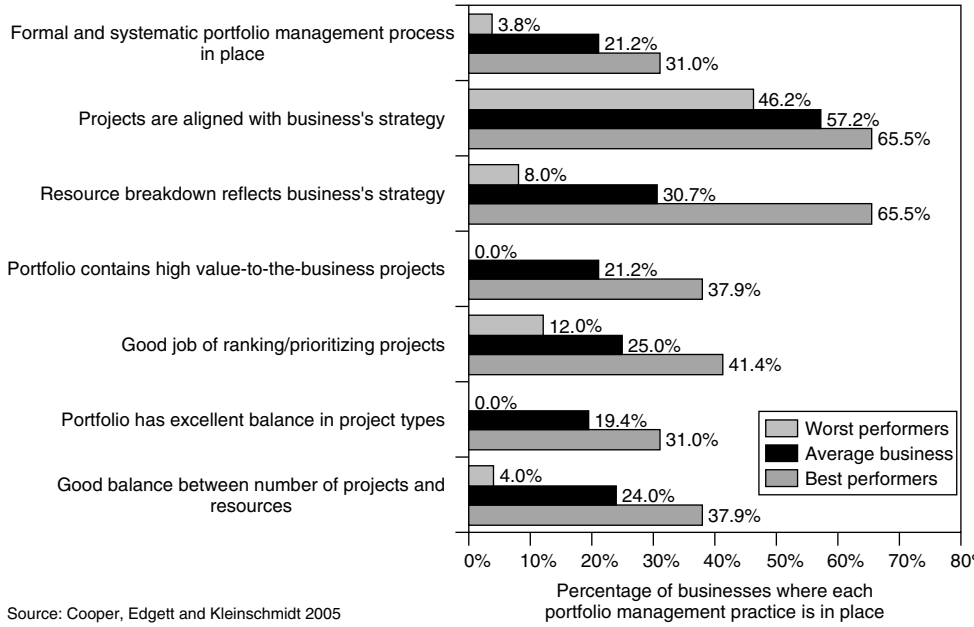


Figure 2 Portfolio management practices – impact on performance (reproduced from Cooper, R.G., Edgett, S.J., and Kleinschmidt, E.J. (2005) Best Practices in Product Development: What Distinguishes Top Performers, Product Development Institute, at www.prod-dev.com for detailed examples and additional supporting data).

management approaches. They have achieved strategic alignment of projects with their innovation strategies. They have prioritized and ranked their projects effectively for value maximization. And, they achieve the right balance and mix of projects.

Research shows that having a portfolio management approach in place seems to be more important than the details of which tools and metrics one elects to use. In other words, any portfolio system is better than no system at all. Businesses featuring systematic portfolio management processes – regardless of the specific approach – outperform the rest (Cooper, Edgett, and Kleinschmidt, 2005; Cooper and Edgett, 2006, 2008).

Using Figure 2, it is possible to measure how an organization's portfolio management methods compare with the top-performing companies. Score the company on each of the seven variables in Figure 2 on a 0–100 scale, where a score of 100 is highest. The questions on which one scores high or low will provide an excellent starting point for assessing and improving the company's portfolio performance.

Additional resources (articles and data) on portfolio management are available for download to the reader at www.prod-dev.com.

Bibliography

Cooper, R.G. (2004) *Product Leadership: Pathways to Profitable Innovation*, 2nd edn, Basic Books, Cambridge.

Cooper, R.G. (2005) Portfolio management for product innovation, in *Project Portfolio Management: A Practical Guide to Selecting Projects, Managing Portfolios, and Maximizing Benefits* (ed. H. Levine), Jossey-Bass, San Francisco.

Cooper, R.G. and Edgett, S.J. (2003) Overcoming the crunch in resources for new product development. *Research-Technology Management*, **46** (3), 48–58.

Cooper, R.G. and Edgett, S.J. (2006) Ten ways to make better portfolio and project selection decisions. *Visions*, **XXX** (3), 11–15.

Cooper, R.G. and Edgett, S.J. (2008) Maximizing productivity in product innovation. *Research-Technology Management*, **51** (2), 47–58.

Cooper, R.G., Edgett, S.J., and Kleinschmidt, E.J. (1997a) Portfolio management in new product

development: lessons from the leaders – part I. *Research-Technology Management*, **40** (5), 16–28.

Cooper, R.G., Edgett, S.J., and Kleinschmidt, E.J. (1997b) Portfolio management in new product development: lessons from the leaders – part II. *Research-Technology Management*, **40** (6), 43–52.

Cooper, R.G., Edgett, S.J., and Kleinschmidt, E.J. (2000) New problems, new solutions: making portfolio management more effective. *Research-Technology Management*, **43** (2), 18–33.

Cooper, R.G., Edgett, S.J., and Kleinschmidt, E.J. (2002a) *Portfolio Management for New Products*, 2nd edn, Perseus Books, Reading.

Cooper, R.G., Edgett, S.J., and Kleinschmidt, E.J. (2002b) Portfolio management: Fundamental to newproduct success, in *The PDMA Toolbox for New Product Development* (eds P. Beliveau A. Griffin, and S. Somermeyer), John Wiley & Sons, Inc., New York, pp. 331–364.

Cooper, R.G., Edgett, S.J., and Kleinschmidt, E.J. (2004) Benchmarking best NPD practices – II: strategy, resource allocation and portfolio management practices. *Research-Technology Management*, **47** (2), 50–59.

Cooper, R.G., Edgett, S.J., and Kleinschmidt, E.J. (2005) *Best Practices in Product Development: What Distinguishes Top Performers*, Product Development Institute, at www.prod-dev.com, for detailed examples and additional supporting data.

launch strategies

Katrin Talke and Erik Jan Hultink

The Importance of an Effective Launch Strategy

The new-product development (NPD) process is regularly seen as a sequence of stages, which ultimately aims at presenting a new product to its target market and at generating income from sales of the new product (*see also* THE STAGE-GATE IDEA TO LAUNCH SYSTEM). The launch stage encompasses all activities facilitating an efficient diffusion of the new product in the marketplace. These activities have been referred to under the collective terms of *launch strategy*, *market entry*, *new-product launch*, *introduction*, or *market launch* (Hultink *et al*., 1997).

Decisions made during the launch phase are recognized as important drivers of new-product performance (*see* SUCCESS FACTORS FOR NEW-PRODUCT DEVELOPMENT) (Hultink *et al*., 1997). The results of a meta-analysis by Henard and Szymanski (2001) underscore that "proficiency of launch" is one of the dominant factors explaining new-product performance. At the same time, new-product launch is often the single most costly step in the NPD process (Hultink *et al*., 1997). Depending on the product and industry, this can include costs for mass-producing the product, setting up the supply chain, establishing dealer and service networks, training the sales force, and advertising the new product. The launch phase also involves large risks as this stage is often the make-or-break point for a product's life cycle. Launch managers are confronted with questions such as whether there will be a market apart from the initial innovators and technology enthusiasts, how competitors will react to the new product, and how competition will evolve over time, and whether the product functions will prove stable, and the technology will become the dominant design. Owing not only to the importance of the launch stage but also to the risks and significant commitments of time, money, and resources, formulating an effective launch strategy is a top priority for firms.

The Anatomy of a Launch Strategy

A review of the new-product launch research of the last 25 years reveals about two dozen studies that seek to illustrate the nature and performance effects of launch activities. These studies use widely varying conceptualizations and operationalizations of new-product launch activities. While earlier studies do not distinguish between strategic and tactical launch decisions (Yoon and Lilien, 1985), more recent approaches regularly propose these two categories of launch decisions (Hultink *et al*., 1997; Di Benedetto, 1999; Langerak, Hultink, and Robben, 2004; Hsieh, Tsai, and Hultink, 2006).

While strategic launch decisions are generally made in early stages of the NPD project, tactical launch decisions occur at later stages, mostly after conceptual and physical development of the new product is complete (Hultink *et al*., 1997). Debruyne *et al*. (2002, p. 161) describe strategic launch decisions as differing from tactical launch decisions in the sense that "they are more important, involve a substantive resource commitment, and are difficult to alter once a trajectory has been selected."

Most studies in this field define strategic launch decisions as comprising of those decisions that set the parameters within which the new product will compete; that is, they define the objectives of the launch, select the market(s) into which the new product will be introduced, determine the competitive position of the new product, and time the entry (Hultink *et al*., 1997). Several authors approach the issue of strategic launch decisions by focusing on the difficulties firms face when launching new products. The choice of the market segment(s) to target, the PRODUCT POSITIONING, and the market entry timing are among the most critical and influential decisions.

Tactical launch decisions can still be modified at a later stage in the launch process (Hultink *et al*., 1997), and are mostly defined along the traditional elements of the marketing mix. Hence, they encompass product, promotion, distribution, and pricing decisions, such as the breadth of product versions launched, the nature and scope of services offered alongside the new product, how and where to distribute

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

and promote the product, and its price (Hultink *et al.*, 1997).

The Impact of Launch Strategies on New-Product Performance

Most studies dealing with the performance impact of launch strategies start off by illustrating the nature of launch decisions and identify different sets or typologies of associations between strategic and tactical launch decisions (Hultink and Schoormans, 1995; Hultink *et al.*, 1997). In a second step, they analyze the relationships between those sets of launch decisions and new-product performance. Several studies also consider contingencies, like the nature of the new product launched (e.g., product innovativeness, product category; *see* INNOVATION TYPOLOGIES) or the market context (e.g., market and technology dynamics).

A brief review of selected studies shows that Hultink and Schoormans (1995) identify two distinctive launch clusters. The "skimming" cluster is characterized by a technology leadership positioning, a pull strategy, a small product assortment, a company brand name, and a skimming pricing strategy. This launch strategy is recommended for an early entry, while for a late entry, the "penetration" cluster is superior. This group is characterized by a small product assortment, a pull strategy, and a penetration pricing strategy.

Manu and Sriram (1996) identify four launch strategy types. "Product innovators" have the highest sales force expenses, invest most in distribution, have the best image and product quality, and offer the best services, followed by the "process innovators" and "late entrant noninnovators." While "product innovators" realize the highest market share gains, "former pioneers" are performing best financially and get the highest relative market share (*see also* FIRST-MOVER ADVANTAGE). The results further show that "product innovators" and "process innovators" are mostly found in dynamic markets.

Hultink *et al.* (1997) identify four launch strategy types for industrial products. "Niche innovators" are innovative products launched by pioneers, which serve market niches with a broad product line, a skimming pricing strategy accompanied by high distribution intensity. "Niche followers" follow with a smaller product line. "Mass marketers" and "would-be me-too" launch strategies adapt new products for the mass market. The most successful launch types are the "niche innovators" followed by the "niche followers" and the "mass marketers."

When contrasting launch decisions for consumer and industrial products in a later study, Hultink *et al.* (2000) report different key success factors for both groups. Successful consumer goods tend to be brand extensions, more innovative, having a similar price level as the competition and show higher promotional and distribution expenditure compared to their competitors. Successful industrial goods are also more innovative, but show a broader product assortment, pursue more activities of direct marketing, and have similar promotional and distribution expenditure compared to their competitors.

From these studies, the conclusion can be drawn that the effectiveness of different launch strategy types depends on product- and market-related contingencies. However, from an aggregate perspective, the findings support the notion that launch performance usually requires setting clear, well-founded launch objectives, defining a precise target market, and positioning the product in a distinct manner. Significant investments in promotional and distribution expenditures also seem to drive launch performance.

Antecedents of Launch Decisions

The question as to which background factors influence decisions made during new-product launch has only rarely been addressed in the new-product launch literature. So far, only few studies have investigated antecedents of an effective launch strategy. Hsieh, Tsai, and Hultink (2006) investigated the impact of a firm's resource configurations on the launch strategies pursued. While strategic and organizational abilities and physical assets seem to be unconnected with an effective launch strategy, the authors find that technological capabilities and societal-network assets and goodwill are crucial resources for companies following a "niche innovators" strategy, while for "mass marketers" only technological capabilities are

relevant. However, pinpointing why firms adopt different launch strategies is only a first step when aiming to identify the antecedents that drive performance through appropriate launch decisions. The link to market performance is investigated by Langerak, Hultink, and Robben (2004) and Talke and Hultink (2009). Langerak, Hultink, and Robben (2004) explored the impact of a market-oriented firm culture on the nature of launch decisions and on new-product performance. While the results show that a market orientation is related positively to the proficiency in market testing, launch budgeting, and the strategic and tactical launch decisions, the mediated model is only supported for the launch tactics.

Talke and Hultink (2009) propose that the capability to successfully launch new products is based on the interplay between the corporate mind-set and the firm's strategic launch decisions. It is argued that performance in launching a new product requires an analytical, risk taking, and aggressive mind-set to set clear and ambitious launch objectives and to proficiently make use of market segmentation and positioning. The results show that the dimensions of the corporate mind-set have a significant impact on most strategic launch decisions, which in turn significantly contribute to market performance. It is found that an analytical posture is driving the setting of clear launch objectives, the selection of target markets, and the positioning of the new product. Risk taking has a significant impact on market segmentation and product positioning, while aggressiveness drives the setting of strategic launch objectives. These findings confirm the importance of investigating antecedents for a successful new-product launch, as the corporate mind-set and also partly the market orientation can serve as background resources that set the framework for successful launch decisions. Despite this, the question as to which antecedents drive a successful launch strategy requires more investigation.

The Management of Diffusion Barriers

Stakeholder opposition as barriers to effective launch decisions and diffusion. Although not explicitly stated in any of the studies reviewed above, diffusion theory (Rogers, 2003) is the underlying theoretical basis of these approaches to conceptualize an effective launch strategy (*see* DIFFUSION OF INNOVATION). Consequently, the focus is on the potential customers and their adoption decisions. However, such a perspective neglects other stakeholders who can significantly influence the performance of any new product's launch and diffusion, like internal stakeholders, and noncustomer stakeholders from the external market and the broader firm environment, such as suppliers, competitors, and other interest groups. A stakeholder perspective seems to allow a more thorough picture of an effective launch strategy.

In an early conceptual work on creating a marketing hype, Wind and Mahajan (1987) argue for targeted interaction with relevant stakeholders like dealers, suppliers, diffusion agents, and the general public to create a favorable atmosphere before the launch of a new product. According to these authors, the key notion is to sufficiently understand the needs of all relevant stakeholders and address them in an adequate form. However, just recently, Talke and Hultink (2010) detailed the barriers that may arise if relevant stakeholders are not adequately considered and the concrete launch activities that have to be pursued to overcome such barriers.

From an internal perspective, employees involved in new-product launch, like marketing, sales, or service personnel, may expose knowledge and motivation-related barriers opposed to launch objectives (Hultink and Atuahene-Gima, 2000). Negative consequences resulting from motivation barriers may manifest themselves in boycotting activities, like omitting critical launch activities or a purposeful ill-doing of launch tactics. Additionally, frontline personnel may experience a lack of knowledge with respect to the new product to be launched. Such a knowledge gap can lead to activities not matching the launch objectives. An inadequate sales approach toward customers or incorrect declarations vis-à-vis other relevant market partners may hamper launch success (Hultink and Atuahene-Gima, 2000).

Talke and Hultink (2010) propose that from the perspective of the external market, one has to differentiate between market parties generally interested in constructive interaction

and those who intend to compete. Market actors interested in cooperation, like suppliers or complementary producers, face various uncertainties. The major worry is concerning the general strategic and financial worth of the investments that are made by the innovating firm during any cooperation. Information about new-product characteristics and benefits (*see* PRODUCT SPECIFICATIONS) are needed to anticipate market acceptance and to fit the own offering. Moreover, uncertainties exist about the innovating firm's expectations concerning the cooperation modalities (Hallikas, Virolainen, and Tuominen, 2002). If these uncertainties and information deficits cannot be reduced by the innovating firm, the cooperation will not happen or work out as expected. The resulting lack of support constitutes a diffusion barrier, which can impede a successful new-product launch in different ways (Debruyne *et al.*, 2002). A lack of availability of suppliers' components in due time and with good quality will jeopardize the new product's availability in the market (Hallikas, Virolainen, and Tuominen, 2002). Without compatible, value-adding offers of complementary producers, the attractiveness of the new product will decrease and the innovation may not get established as a standard. Without the support of sales agents, the new product's distribution may be endangered; and a lack of service offerings, as consulting, training, or installation by adequate service providers, can impede customer adoption.

Market players intending to compete will try to build diffusion barriers in favor of their own product (Debruyne *et al.*, 2002). Measures may include lowering prices, increasing communication expenditures, modifying the existing product offering, imitating the innovation, or trying to establish another standard. Debruyne *et al.* (2002) find that the characteristics of the launch strategy have a significant impact on both the occurrence and nature of competitive reactions. Competitors fail to respond to radical innovations and to new products that employ a niche strategy. However, they do react to new products that can be assessed within an existing product category and thus represent an unambiguous attack. Both innovative and imitative new products meet with reactions in this case. Competitors are more inclined to react to the introduction of new products that are supported by extensive communication by the innovating firm. The likelihood of reaction is also higher in high-growth than in low-growth markets. For the innovating firm, the consequences of diffusion barriers set by competitors can extend from reduction of market share to market extrusion.

Stakeholders of the broader firm environment, such as the general public, legal and political institutions, or other interest groups, also face uncertainties about the consequences of the new product and how it fits their own goals. A new-product offering that offends values of interest groups can lead to complaints, protests, or boycotts against the new product. This may result in image loss or legal regulations that can significantly hinder a new product's diffusion (*see* DIFFUSION OF INNOVATION). Overall, opposition of different stakeholder groups can constitute a threat to launch success. Being aware of such potential diffusion barriers is, however, only a first step preceding an active management of such barriers.

Managing diffusion barriers. When comparing the impact of addressing these different diffusion barriers on launch success, Talke and Hultink (2010) find that first of all, barriers related to customers, suppliers, and stakeholders of the firm environment need to be lowered. The activities they propose are on an aggregate level and concern the degree to which the innovating firm interacts with these parties; that is, explicitly addresses them with targeted launch communication or cooperates with them in the preparation of or during a new-product launch. For the group of competitors, Talke and Hultink (2010) find that a balanced launch approach including measures to both lowering and erecting entry and diffusion barriers increases the market performance of new products. Under high uncertainty (i.e., in the case of complex products, and high market and technology dynamics) managing multiple diffusion barriers is of higher relevance than in more unambiguous, clear-cut contexts.

Selectively, authors active in the field of launch strategy have dealt with activities addressing one specific stakeholder group. Exemplary studies with a focus on launch

communication and on managing the firm's frontline personnel will be reviewed in further detail below.

Managing diffusion barriers with launch communication. Several authors view decisions around communication activities as critical to a successful launch as these decisions are most directly responsible for aiding the market acceptance of a new product (Lee and O'Connor, 2003). In a recent article, Michael and Palandjian (2004) confirm that performance in new-product launch increases with spending on launch communication. Apart from the amount of investment in launch communication (Michael and Palandjian, 2004), other studies focus on issues such as the impact of characteristics of communication channels (Hultink *et al.*, 2000), type of communication strategies (Schatzel and Calantone, 2006), or communication content (Talke and O'Connor, 2009) on new-product performance.

A particular communication strategy relevant in the launch context is preannouncement. Preannouncements can be understood as prelaunch communication activities directed at market players including customers, distributors, suppliers, investors, and journalists with the intent to reduce uncertainty in the firm's favor. Preannouncements provide market partners interested in cooperating with the opportunity to better plan, thus deterring a competitor's entry into a market segment, and engendering positive opinions from market players regarding the firm and its intentions (Schatzel and Calantone, 2006). Schatzel and Calantone (2006) found that the more a firm engages in preannouncements, the greater is the favorable bias and interest of various market parties. Lee and O'Connor (2003) examine the effect of different preannouncement strategies (targeting at customer education, anticipation creation, and market preemption) on new-product performance. They find that preannouncements emphasizing customer education lead to higher performance for products that are difficult to adopt. On the other hand, a preannouncement message that focuses on creating anticipation for the new product leads to higher performance for products that provide high superiority to customers.

Concerning launch communication content, Talke and O'Connor (2009) investigate the specific content that facilitates new-product diffusion. They propose a model of three message content dimensions: usability information, technical information, and financial information. Each of these information types is designed to mitigate a set of perceived risks that consumers may experience in the decision to adopt a new industrial product. They find that usability information and financial information are effective in improving market performance, while technical information is countereffective. When launching highly innovative products, usability information is of specific importance, while financial details become less important. For pioneers, communicating the new product's usability is highly effective, while for later entrants using financial information increases performance.

Managing internal diffusion barriers. The firm's frontline personnel can be seen as the first stakeholder group that needs to be won over to optimize launch performance. Concerning the question of how to effectively manage the firm's sales force, Hultink and Atuahene-Gima (2000) find that in a situation of increased uncertainty, salespeople perform better when their sales manager provides them with feedback on their past performance. As feedback reminds salespeople of the results that they are expected to attain, they may feel less ambivalent about their role, and will be better armed to experiment and take risks in selling the new product. This finding suggests that if salespeople are rewarded on the basis of their outputs, they perform better.

Conclusion

Given the widely acknowledged importance of the launch phase, one may expect a thorough understanding of critical launch activities, a sound theoretical basis, and a body of empirical findings substantiating measurement models, performance relationships, contingencies, and antecedents. While in recent years there has been an increasing interest in the last phase of the NPD process, there is still a lack of theory-based launch research. As a consequence,

while much progress has recently been made, there is still ample room to further investigate the activities that a firm should pursue around a new-product launch, resulting in a demand for further research in the areas of conceptualizing launch activities and analyzing related effects.

Bibliography

Debruyne, M., Moenaert, R.K., Griffin, A. *et al*. (2002) The impact of new product launch strategies on competitive reaction in industrial markets. *Journal of Product Innovation Management*, **19**, 159–170.

Di Benedetto, C.A. (1999) Identifying the key success factors in new product launch. *Journal of Product Innovation Management*, **16** (6), 530–544.

Hallikas, J., Virolainen, V.M., and Tuominen, M. (2002) Understanding risk and uncertainty in supplier networks – a transaction cost approach. *International Journal of Production Research*, **40** (15), 3519–3531.

Henard, D.H. and Szymanski, D.M. (2001) Why some new products are more successful than others. *Journal of Marketing Research*, **38**, 362–375.

Hsieh, M.-H., Tsai, K.-H., and Hultink, E.J. (2006) Relationships between resource configurations and launch strategies in Taiwan's IC design industry: an exploratory study. *Journal of Product Innovation Management*, **23** (3), 259–273.

Hultink, E.J. and Atuahene-Gima, K. (2000) The effect of sales force adoption on new product selling performance. *Journal of Product Innovation Management*, **17** (6), 435–450.

Hultink, E.J., Griffin, A., Hart, S., and Robben, H.S. (1997) Industrial new product launch strategies and product development performance. *Journal of Product Innovation Management*, **14** (4), 243–257.

Hultink, E.J., Hart, S., Robben, H.S., and Griffin, A. (2000) Launch decisions and new product success: an empirical comparison of consumer and industrial products. *Journal of Product Innovation Management*, **17**, 5–23.

Hultink, E.J. and Schoormans, J.P.L. (1995) How to launch a high-tech product successfully: an analysis of marketing managers' strategy choices. *Journal of High Technology Management Research*, **6** (2), 229–242.

Langerak, F., Hultink, E.J., and Robben, H.S. (2004) The impact of market orientation, product advantage, and launch proficiency on new product performance and organizational performance. *Journal of Product Innovation Management*, **21**, 79–94.

Lee, Y. and O'Connor, G.C. (2003) The impact of communication strategy on launching new products: the moderating role of product innovativeness. *Journal of Product Innovation Management*, **20**, 4–21.

Manu, F.A. and Sriram, V. (1996) Innovation, marketing strategy, environment, and performance. *Journal of Business Research*, **35**, 79–91.

Michael, S.C. and Palandjian, T.P. (2004) Organizational learning and new product introductions. *Journal of Product Innovation Management*, **21**, 268–276.

Rogers, E.M. (2003) *Diffusion of Innovations*, New York: Free Press.

Schatzel, K. and Calantone, R. (2006) Creating market anticipation: an exploratory examination of the effect of preannouncement behavior on a new product's launch. *Journal of the Academy of Marketing Science*, **34** (3), 357–366.

Talke, K. and Hultink, E.J. (2009) The impact of the corporate mindset on new product launch strategy and market performance. *Journal of Product Innovation Management*, in print.

Talke, K. and Hultink, E.J. (2010) Managing diffusion barriers when launching new products. *Journal of Product Innovation Management*, in print.

Talke, K. and O'Connor, G. (2010) Conveying effective message content when launching new products. *Journal of Product Innovation Management*, in print.

Wind, Y. and Mahajan, V. (1987) Marketing hype: a new perspective for new product research and introduction. *Journal of Product Innovation Management*, **4**, 43–49.

Yoon, E. and Lilien, G.L. (1985) New industrial product performance: the effect of market characteristics and strategy. *Journal of Product Innovation Management*, **3**, 134–144.

growth strategies

Michael J. Baker

Introduction

In a competitive world, the status quo is in a state of constant flux. Standing still is not an option. Only through a process of constant adaptation and renewal can one hope to survive and prosper. It is for this reason that organizations invest in new-product and service development, which is the theme of this volume of the encyclopedia. But, for investment to pay off, it has to add value and generate sales in excess of its costs. And to achieve this, one needs a growth strategy.

From the firm's perspective it is important to distinguish between macroeconomic and microeconomic conditions. Understandably, a firm's primary preoccupation is with the products and services that it offers and the markets that it serves. This is the domain of microeconomics and a firm's performance (growth or otherwise) is determined by its success in "jockeying for position" in competition with other firms serving the same market. However, microeconomic conditions are ultimately subject to the prevailing macroeconomic conditions. Clearly, if the global economy is expanding, growth opportunities are much greater as consumers have more discretionary purchasing power with which to buy goods and services. For individual firms, taking advantage of this is still a challenge. But, at the time at which this is being written, the global economy has suffered a major economic downturn and has moved into recession. In a recession, overall effective demand contracts (growth is negative) but the impact of this will not affect all firms and markets in the same way, and the challenge becomes acute as does the selection and implementation of a strategy.

To address this challenge, we begin with a discussion of strategy and the imperative of growth. Next, we review the growth-vector matrix as a diagnostic framework for identifying alternative strategies available to the firm. These strategies depend largely on the stage at which the industry and its products are in their life cycles – the theme of the next section. The ideas of the growth-vector matrix and product life cycles (PLCs) are then combined in the diagnostic developed by the Boston Consulting Group, which is reviewed as a basis for examining the options available to the firm at each of its major phases.

The Imperative of Growth

While a multidivisional organization might require a "corporate" strategy to coordinate and direct the separate strategies of the strategic business units (SBUs) that comprise it, the great majority of business organizations are SBUs in their own right. When this is the case, then the marketing strategy is both a competitive and a corporate strategy.

According to Arthur D Little,

> A strategic business unit—or strategy centre—is a business area with an external marketplace for goods and services, for which management can determine objectives and execute strategies independent of other business areas. It is a business that could probably stand alone if divested. Strategic business units are the "natural" or homogeneous business of a corporation.

Central to the SBU is the *product*, which we define here as a tangible offer the firm makes to the market and which includes quality, range, design, and branding. It also includes services such as delivery, leasing, repair, and training (Doyle, 2000). This is what the producer or supplier offers to a potential customer in exchange for something else (usually money), which the supplier perceives as equivalent to or of greater value. It follows that product strategy lies at the very heart of corporate and marketing strategy.

In broad terms, an organization has only a limited number of strategic options available to it. Leaving aside the possibility of withdrawing from a market, the focus of these options is growth – the subject of this article. According to the AMA American Marketing Association dictionary entry for "growth strategy":

> Market share expansion is the prime objective under this strategy, even at the expense of short-term earnings. The firm may seek to expand market share through a number of alternative routes. First, the firm may seek new users who may previously have been loyal to

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

other brands, or tended to switch, or were not users of the category at all. The second way in which the firm can expand its market share is to expand usage by current users; for instance, by identifying and promoting new uses.

This is a very narrow view that is restricted primarily to a firm with a share of a mature market for a product or service and largely ignores the need for a firm to continually update its offerings if it is to remain competitive (*see* MANAGING MATURE PRODUCTS).

In "Value-based Marketing," the late Prof. Doyle (2000) devotes a complete chapter to what he calls "*the growth imperative*." In his view, management should make growth a top priority as it is essential for creating shareholder value because of its potential for increasing the net present value of future cash flows. Despite this, many managements prefer to emphasize rationalization of the present business because not only is it easier and less risky, it also has a more immediate effect on increased profits and cash flow. However, such effects may be purely short term and so threaten the future survival of the company. Of course, by then, the management responsible may well have retired! By contrast, a growth strategy is likely to be more risky in that it involves seeking to anticipate what customers will want in the future and adjusting the firm's product market mix to anticipate these future changes. This is not to discount the important contribution that rationalization may make in terms of improving both efficiency and effectiveness. But, for long-term survival a growth strategy is essential.

To succeed, a growth strategy requires a firm foundation from which to develop. Invariably, this firm foundation is to be found in the trust and loyalty of the firm's existing customer base as success in the short term depends very much upon repeat purchase and word-of-mouth recommendation. These two activities comprise the first two steps of nine in what Doyle calls "*the ladder of growth*." These nine steps are defined as follows:

1. increasing customer retention rate
2. increasing customer share
3. gaining new customers
4. developing new products and services
5. entering new markets
6. exploring new distribution channels
7. increasing international growth
8. negotiating acquisition and alliances
9. exploring growth outside current industry boundaries.

Each of these steps is then described in some detail with an explicit discussion of the tactics and activities associated with them. Given the space limitations for this article, a simpler conceptualization consisting of fewer steps seems appropriate. Such a conceptualization is to be found in the growth-vector matrix first proposed by Ansoff (1957). (In passing it should be noted that earlier in the same year Johnson and Jones (1957) had published a very similar "technology market matrix." However, this has attracted relatively little attention, probably because this matrix too contains nine options or alternative strategies, whereas Ansoff's matrix contains only four.)

The Growth-Vector Matrix

The "growth-vector matrix" is a familiar analytical device to be found in most textbooks dealing with marketing strategy and is closely associated with the name of Igor Ansoff. The original version first appeared in the September/October 1957 issue of the Harvard Business Review in an article entitled "Strategies for diversification" but was subsequently revised in his book on corporate strategy published in 1965. In Figure 1, we offer an adaptation of his original idea, which retains a two-by-two format but conceptualizes it as reflecting the interaction between demand and supply, as represented by customers and technology/products, both current and "new". This results in four basic alternatives, which Ansoff described as market penetration, market development, product development, and diversification. The following are defining characteristics of these alternatives:

1. *Market penetration*: The company seeks increased sales for its present products in its present markets through more aggressive pricing, promotion, and distribution. By these means, the firm seeks to increase the volume of consumption by its existing

Supply / Demand	Current Technology / Products	New Technology / Products
Current Customers	Market penetration	Product development
New Customers	Market development	Diversification

Figure 1 Growth-vector components (adapted from an original idea by Ansoff (1965)).

customers, convert nonusers to users, and attract exiting users from its competitors.

2. *Market development*: The company seeks increased sales by taking its present products into new markets. This may be achieved by expanding into new geographical areas, for example, exporting, and/or by opening up new distribution channels.
3. *Product development*: The company seeks increased sales by developing improved products for its present markets. In doing so the firm is seeking to achieve the same objectives as a penetration strategy with the added potential for entering new markets resulting in diversification.
4. *Diversification*: The company seeks increased sales by developing new products and new markets. Where the firm attempts to do this without prior knowledge/experience of either the product (technology) or market (customer) the risks are far greater than when it "migrates" into this state via a strategy of product or market development.

If we compare this classification with Doyle's growth ladder, we can see that market penetration is equivalent to steps 1–3, market development embraces steps 5–7, while step 4 is equivalent to product development. Step 8, acquisitions and alliances, may relate to the current product–market mix or may be addressed at step 9, which is equivalent to diversification. As can be seen from this figure, the AMA definition is focused primarily on what Ansoff terms *market penetration* and, to a lesser degree, *market development*. As such, it ignores both new-product development and diversification, which are major growth strategies.

A more appropriate and dynamic approach is to consider the management of a product as it progresses through its life cycle and recognize that a different emphasis and marketing mix is called for in each of the major phases – introduction, growth, maturity, and decline. Thinking on these issues was strongly influenced by the work of Theodore Levitt.

Growth and Product Life Cycles

Theodore Levitt was one of the most influential marketing thinkers of the second half of the twentieth century (*see also* PRODUCT LIFE CYCLE). Central to his writings is a concern with the marketing concept and competition at the industry/firm level. Many would regard his "marketing myopia" as the catalyst that heralded the adoption of a marketing orientation as a key to competitive success. In the opening paragraph, he asserts as follows:

> Every major industry was once a growth industry. But some that are now riding a wave of growth enthusiasm are very much in the shadow of decline. Others which are thought of as seasoned growth industries have actually stopped growing. In every case the reason growth is threatened, slowed, or stopped is not because the market is saturated. It is because there has been a failure of management.

In a nutshell, his argument was that management had failed to adapt to new and improved ways of satisfying customers by emphasizing the products and services they offered for sale rather than the needs that they served. Underpinning this argument was the view that technological innovation offered the opportunity to create new and better products that would both substitute for existing products and stimulate new demand and growth. In other words, technologies and the products derived from them have a life cycle each phase of which called for different marketing mixes and managerial action.

While some have questioned the value of the PLC concept, there can be no doubt that it offers a useful description of the phases through which a successful products pass from their initial introduction to their eventual withdrawal from the market place. Once the characteristics of these phases have been identified, it becomes possible to recommend courses of action best suited to their operational management as illustrated in Table 1.

The Boston "Box"

Although rarely acknowledged explicitly, the notions of the growth-vector matrix and the PLC are combined in what is commonly referred to as "*the Boston Box*" (*see also* GE/MCKINSEY MATRIX).

In discussing product policy and strategy, and especially new products, the focus is frequently on a single product. In reality, most firms have a range or portfolio of products each of which may be at a different stage in its life cycle. It is recognition of this that led the Boston

Table 1 Product life cycle: implications for marketing.

Product Life Cycle	*Introduction*	*Growth*	*Maturity*	*Decline*
Characteristics				
Sales	Low	Fast growth	Slow growth	Decline
Profit	Negative	Rapid rise	Falling margins	Declining
Cash flow	Negative	Moderate	High	Moderate
Strategy				
Objective	Aggressive entry	Maximize share	Boost profits	Milk product
Focus	Nonusers	New segments	Defend share	Cut costs
Customer targets	Innovators	Early adopters	Majority	Laggards
Competitor targets	Few, preempt	Growing number	Many declining	
Differential advantage	Product performance	Brand image	Price and service	Price
Marketing Mix				
Product	Basic	Extensions and enhancements	Differentiation, variety	Rationalize range
Price	High	Lower	Low	Stabilizing
Promotion	High	High	Falling	Low
Advertising forms	Awareness	Brand performance	Loyalty	Selective
Distribution	Selective	Intensive	Intensive	Rationalize
Organization				
Structure	Team	Market focus	Functional	Lean
Focus	Innovation	Marketing	Efficiency	Cost reduction
Culture	Freewheeling	Marketing led	Professional	Pressured

Source: Reproduced with permission from Peter Doyle (1999), Product Life Cycle Management, in IEBM Encyclopedia of Marketing, Ed. Michael J Baker, London: Thomson.

Consulting Group to propose a diagnostic procedure analyzing the comparative and collective performance of different products in a firm's range. It was this that was to become famous as the "Boston Box." While several variants of the Boston Box have been proposed, they all rest on two fundamental propositions, namely, the competitive value of market share depends on the structure of competition and the stage of the PLC.

The importance of taking into account the stage that a product (or industry) has reached in its life cycle is fundamental to the whole concept and practice of marketing. As the marketing concept makes clear, survival depends upon understanding and satisfying the needs of customers, and realization and acceptance of the fact that, over time, these needs will change, often owing to the action of marketers; hence the importance of the product life cycle. If, then, we accept the inevitability of the PLC it follows that, even where we are concerned with a basic product such as steel with a life span of decades, we would be unwise to assume that any particular formulation or, more important, method of manufacture is immune from competition and replacement. To guard against this possibility, product differentiation has become a key competitive strategy and new-product development an essential activity, so that even the overtly single-product firm should have a range of variant products at different stages of development.

Like the growth-vector matrix, the Boston Box uses two parameters – market growth, which is a proxy for the stage in the PLC, and market share, which is a proxy for profitability. Its purpose is to distinguish the four basic categories of product, the desired movement of businesses over time, and the likely movements of cash between them. A version of this matrix is presented as Figure 2.

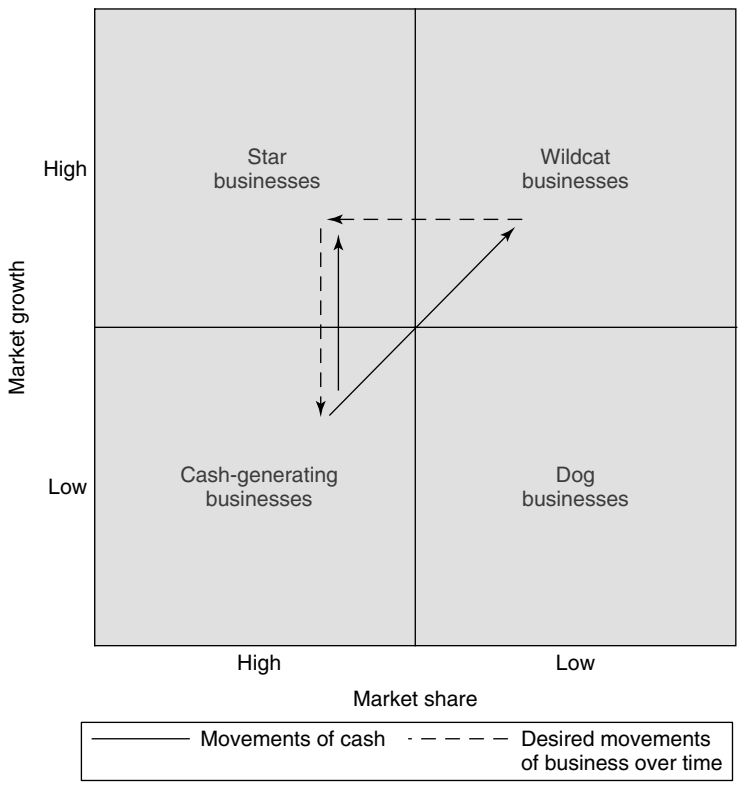


Figure 2 The Boston box (reproduced with permission from Conference Board, The (1974) 'Why Products Fail').

Wildcat businesses (also known as *problem children*) represent products at the beginning of their life cycle where hopes run high, and expectations of rapid growth and profitability encourage the firm to invest heavily in its new product despite uncertainty as to its eventual success. Clearly, the hope is that by doing so the product or business will get onto the rapid-growth curve when sales expand exponentially, and costs and overheads decline rapidly because of economies of scale and growing experience. At this stage of market development, competition is fierce as firms jockey for position to become "stars," and seek to secure cost and price advantages, which will enable them to secure a dominant share of the new market.

Cash cows represent products that have survived the shakeout of the growth phase when firms compete aggressively with one another and have become established players in a mature market. To displace these dominant products, new and better alternatives are required and a new cycle of innovation occurs. As new wildcat products become established and grow, demand for mature products begins to fall and move into decline in terms of their market share. However, it is important to recognize that the decline in market share does not necessarily equate with a decline in growth. For example, in the packaging industry during the 1950s, new synthetic plastic packaging materials began to substitute for the more traditional packaging materials such as paper, board, and tinplate. Today the latter materials comprise a much smaller share of the total market for packaging materials but, owing to overall growth of demand for packaged goods, the demand for paper, board, and tinplate today is far higher than it was 50 years ago.

Product Differentiation as a Growth Strategy

On the basis of their visits to a number of leading companies, Deschamps and Ranganath Nayak (1993) take the view that such companies succeed competitively because they believe there is no such thing as a "commodity" product (*see also* PRODUCT DIVERSIFICATION). Even in basic industries such as minerals, chemicals, and agribusiness "suppliers can always find ways to differentiate their products from those of their competitors". The corollary is that if one fails to differentiate, the only basis for competition is cost and price.

Deschamps and Nayak propose five basic strategies for competing through products. These they designate as follows:

1. competing through product proliferation
2. competing through value
3. competing through design
4. competing through innovation
5. competing through service.

Competing through product proliferation. This strategy was pioneered by General Motors in the 1920s and 1930s in its battle with Ford. Once thought to be the preserve of large, rich companies, it is now recognized as reflecting the company's development effectiveness and seen as one of the most effective strategies for all companies. Japanese companies were seen as being particularly adept at proliferation strategy in which they took "a scattershot approach to the marketplace," which is contrasted with Western companies that have a "rifle mentality." Thus Japanese firms will launch many new products simultaneously or in close succession and use market response to determine which to support and which to withdraw – what, elsewhere has been called "*trial and error marketing*." By contrast, Western firms seek to define their single shot and its target much more precisely with the result that if they miss they have to repeat the procedure again.

Key exponents of a proliferation strategy are Honda in motorcycles who overwhelmed Yamaha's challenge in the early 1980s with 80 new models and 113 product offerings against Yamaha's 34 and 37 in an 18-month period; Sony, whose Walkman technology could not be protected so that 80% of the market went to competitors until Sony introduced 150 new models in the period 1981–1989 and regained preeminence with over 40% market share; and Procter & Gamble, who introduced numerous variants of its Pampers brand to fight off competition.

Competing through value. Offering high quality at low cost is seen by some to be the ultimate strategy (*see* VALUE PROPOSITION).

"In the hands of daring entrepreneurs such as Henry Ford, this strategy spurred the great consumer revolution of the twentieth century." Deschamps and Nayak identified two discrete approaches – *continuous improvement* and *radical restructuring*

"Many companies in traditional industries operate instinctively within a relatively simple quality/cost trade-off model. In this way of thinking, there is a single, fixed, direct relationship between quality and cost. Any significant improvement in the product, whether in design features, or even quality 'feel', necessitates a proportional increase in costs." Many firms are seen to support this view.

By contrast, Toyota has "a mental model of multiple and movable trade-off curves. This alternative view allows for the systematic planning and introduction of process improvements that dramatically shift any given quality/cost curve." Shifting from one quality/cost curve to another is achieved by a number of complementary approaches including process efficiencies in development and manufacturing, systematic design improvements, and "a persistent tracking of inefficiencies and waste at all levels."

The merits of these alternative approaches have been widely recognized and adopted in what Senge (1990) describes as the "*learning organization*." The approach of radical restructuring is typified by IKEA, which "revolutionized furniture design, manufacturing, and retailing by rethinking the entire business system of its industry from product concept to distribution." By offering excellent design and product quality backed up by strong customer service and a liberal return policy for its range of self-assembly furniture, IKEA offers superior value when compared to conventional furniture manufacturers/retailers.

Competing through design. For manufacturers who compete through design excellence, " ... design is not a cosmetic add-on but a means of expressing their corporate identity in the marketplace and establishing their products as synonymous with quality" (*see* PRODUCT DESIGN). This is achieved "by designing products that are

- aesthetically appealing
- safe and pleasing to touch and use
- immediately intelligible and easy to operate
- easy to install, handle, store, clean, and maintain
- easy and economical to manufacture."

Success in design calls for a combination of engineering design – from the inside out – with aesthetic design from the outside in.

Competing through innovation. Few companies compete consistently through innovation. Those that do include Black & Decker, Canon, DuPont, 3M, Merck, Phillips, and Sony.

Because of the risks of innovation, fast followers often reap the rewards lost by the true innovator (*see* FIRST-MOVER ADVANTAGE). "However, fast followers usually succeed only when the original innovator lacks the market position or financial strength to fully exploit its innovation. Few fast followers succeed against a healthy, well positioned innovator."

Innovation strategies may be incremental or breakthrough and may be either top down or bottom up in their origins and processes (*see* RADICAL INNOVATION; INNOVATION TYPOLOGIES). Radical (breakthrough) innovations may be the result of a conscious policy. But, according to Deschamps & Nayak, "top-down breakthroughs happen only when these are all in place:

1. a top management with a strong vision of where and how to innovate and the capability to communicate and mobilize people to make it happen;
2. a strong technological culture and world-class capability to develop innovation enabling technologies and new proprietary product concepts;
3. a very clear sense of the customer (through a combination of research and intuition) and the ability to translate product concepts into attractive, saleable products;
4. an ability to combine mutually reinforcing innovations (for example, in product and in manufacturing process)."

That said, bottom-up breakthroughs probably dominate (if for no other reason than that there are many more of them). However, incremental innovation is more commonplace particularly in

the top-down mode where far more R&D policies exist. For bottom-up incremental innovation to occur the right culture has to exist, which establishes a climate that encourages innovations and provides mechanisms for facilitating them.

Competing through service. While customers regard product and service as "two faces of the same coin" most manufacturers give greater emphasis to the product and see services as adding costs rather than value. This perception is being rapidly eroded with the growing recognition of the importance of customer service.

The authors conclude: "Underlying all these diverse ways to compete is the realization that products are created through a *process*."

Leveraging new-product growth. On introduction to a market, a new product has to overcome at least three kinds of potential resistance before it is likely to "take off" and grow rapidly. These may be characterized as technological, infrastructural, and behavioral. Lele (1992) illustrates their effect by reference to the market for microwave ovens. The first designs of microwave ovens were introduced in 1957 but were large and expensive. The technology needed considerable refinement. Secondly, microwave cooking calls for special containers and most preprepared and convenience foods were not packed in the right materials. The infrastructure had to develop to make microwave meals available. Third, until more women began to enter the workforce there was an absence of a sufficiently powerful precipitating circumstance, convenience, to prompt the behavioral response. Lele describes this as the "emerging" phase during which typically there may be many, often small, participants with a wide variety of products and marketing mixes largely due to the low resistance to entry.

In some senses, we may regard the emergent or introductory phase as a period of trial and error, which will lead to the evolution of a dominant design widely accepted by both producers and users. It is this dominant design that usually triggers the rapid-growth phase and sees a switch in emphasis from product development to process innovation. Sellers and buyers are now agreed on the desirable attributes of the new product and competitive advantage resides in being able to produce and market the new product in the most efficient and effective way. During the growth phase, Lele (p. 201) argues that "the emphasis should be on identifying the most likely and profitable product/market areas for introducing flankers, niche products, or complements."

During the growth phase, competitors may choose to follow one or other of three broad alternative strategies. They can attempt to develop a second standard, can adopt the first standard and compete head-on, or can identify potential niches and concentrate on getting into them first. Developing a second standard has major cost implications and, in effect, is equivalent to launching a new product, which would have to survive launch and introduction while competing with the first standard in its rapid-growth phase. This is not exactly an easy or attractive proposition. Competing head-on also requires considerable resources but is often the approach taken by large firms pursuing the strategy of the fast second who delay entering the market until the onset of rapid growth indicates it is propitious for them to climb on the bandwagon. For smaller players this is not a viable alternative. Most competitors will seek to develop specialist niches or complementary products and so ride on the coat-tails of the industry leader.

Sustaining Differentiation

It is clear that a new product enjoys a competitive advantage because it is differentiated (*see* COMPETITIVE ADVANTAGE). It begins to lose this advantage as soon as similar or improved versions are introduced to the market by other firms. It is generally accepted that only two basic strategies are available – cost leadership and differentiation, although Porter and others identify focus as a special case of differentiation. Cost leadership is a function of size and market share and is, by definition, a strategy limited to one (monopolist) or very few (oligopolists) in any given market. It follows that for the great majority of innovators the only viable strategy is differentiation. With a new technology, further developments and improvements are possible, so product development remains a major source of differentiation during the growth phase, with the other elements of the marketing mix playing a supporting role. However, when maturity sets in further, significant product improvements are

no longer available, and competition revolves entirely around manipulation of the price, place, and promotional variables.

As Levitt pointed out in his famous article "Exploit the product life cycle" (*Harvard Business Review*, Nov–Dec 1965): "The ensuing fight [during the market growth phase] for the consumer's patronage poses to the originating producer an entirely new set of problems. Instead of seeking ways of getting consumers to *try the product*, the originator now faces the more compelling problem of getting them to *prefer his brand*" (p. 33). In part, such preference will depend upon having a better, or equivalent product, when first-mover advantages may lend added value, that is, the company's name is most strongly associated with the new product. But, if the firm's product is seen as being of much of a muchness with all the new competitive offerings then price, promotion, and distribution will assume increased importance in helping to differentiate the product.

There is considerable evidence to suggest that many innovators fail to reap the full reward of the growth phase. One reason for this is that the firm's own growth cycle may not reflect that of the industry as a whole. Indeed, the sudden upsurge in growth engendered by the rush of imitators seeking to cash in on the new market opportunity may stall the innovator's own growth curve. As Levitt (1965) comments, "for the originating company its growth stage now becomes truncated. It has to share the boom with new competitors. Hence the potential rate of acceleration of its own take-off is diminished and, indeed, may actually fail to last as long as the industry's" (p. 86). It is for this reason that monitoring both the product and industry life cycles is so important in devising strategies to maintain the firm's competitive edge. As the kind of analysis described by Levitt makes clear, the impact of competition, which is a major driver of rapid growth, will have negative consequences for the originator, *unless* they heed the warnings of life-cycle analysis and take deliberate steps to stay ahead of the game and control market growth.

A classic example of a firm doing this was Pilkington with the introduction of float glass. This process, patented by them, was a major technological breakthrough which offered a product far superior to that made by traditional methods. As a known need innovation with high relative advantage, high compatibility, low complexity (in use), high divisibility (buy a square meter or a square kilometer), and excellent communicability, it was a surefire winner. Everyone wanted it. Clearly, there was no way Pilkington could cope with such a demand on a global scale. Equally clear was the fact that major competitors like St Gobain and Corning would do everything they could to replicate the process. Rather than encouraging a competitive response Pilkington defused it by offering licenses to major producers in other countries and taking a royalty off every square meter sold.

Armed with the hindsight that PLC analysis offers, innovators can preplan their competitive moves so as to avoid being outmaneuvered and overtaken by the competition. To quote Levitt again, "advance planning should be directed at extending, or stretching out, the life of the product. It is this idea of *planning in advance* of the actual launching of a new product to take specific actions later in its life cycle – actions designed to sustain its growth and profitability – which appears to have great potential as an instrument of long-term product strategy" (p. 88).

Much of Levitt's analysis then focuses on extension or stretching strategies which mainly come into play when a product reaches maturity. However, bearing in mind the earlier comment that the onset of competition and rapid growth may truncate the originator's own growth cycle, it is quite likely that the originator will move into the mature phase of their PLC while the industry life cycle is still growing. This, too, may be a source of competitive advantage as the innovator may be able to consolidate their position by adopting extension strategies while the remaining competitors slug it out for a share of a market, which still appears to be growing because of the aggressive competition.

Although profit margins usually decline through the maturity phase products at this stage of their life cycle almost invariably comprise the backbone of an established firm's business and generate most cash for reinvestment in the future. Firth (1980) cites seven reasons given by Rogers, which account for declining profit margins:

1. There is an increase in the number of competitive products leading to overcapacity and intensive competition.
2. Market leaders are under growing pressure from smaller competitors.
3. There is an increase in R & D to find better versions of the product.
4. Cost economies are used up.
5. There is a decline in product distinctiveness.
6. Declining sales result in dealer apathy and disenchantment with a product.
7. The loyalty of those first to adopt begins to waver with changing market composition.

In turn, and because of this profit erosion, the industry tends to stabilize as a set of well-entrenched competitors all seeking a competitive advantage. This they usually attempt to do by adopting one of four basic strategies, which have been characterized as follows:

1. an offensive or "take-off" strategy
2. a defensive strategy
3. a recycle strategy
4. a stretching and harvesting strategy.

An offensive strategy has as its objective a major extension of the PLC and will often lead to one or more periods of renewed growth followed by stabilization at an overall higher level of sales. Perhaps the best-known example of this strategy is that reported by Levitt for nylon in which four quite distinct approaches were used to extend the life cycle:

1. promoting more frequent usage among current users;
2. promoting more varied usage among current users;
3. attracting new users;
4. developing new uses for the basic material.

In terms of Ansoff's growth-vector matrix, approaches 1 and 2 above are equivalent to market penetration, 3 is equivalent to market development and 4 to new-product development. Research by Baker and Hart (1987) into the sources of competitive success in UK industry confirmed that the most successful firms pursued strategies of market penetration, new-product development, and market development simultaneously – precisely the approach taken by Du Pont with nylon.

Where it is not possible to develop new uses or markets for a product, it becomes necessary to protect one's existing share of the market through a defensive strategy. This is often referred to as *dynamic adaptation* and involves manipulation of the nonproduct elements of the marketing mix – price, distribution, and promotion.

A recycling strategy is most often found in the markets for fast-moving consumer goods (FMCG), but is also applicable to consumer durables and industrial products. Like a defensive strategy, recycling seeks to preserve a product's market share against erosion through a preplanned series of relaunches based upon one or more elements of the mix, for example, product improvements, repackaging, new advertising campaign, different channels, and so on.

Stretching or harvesting strategies are most common for products with high market shares and little or no direct competition. Because of its dominant position such a product requires below-average marketing support on a unit basis and so enjoys above-average profit margins.

Summary

> When the shouting is over, one fact is clear: what differentiates perennially great companies from others is the products they sell (Deschamps and Nayak).

We argued that for long-term survival a growth strategy is essential and followed Doyle (2000) in identifying nine options available for achieving this. These were reduced to the four alternatives suggested by the growth-vector matrix based on an idea proposed by Ansoff (1957), namely, market penetration, new-product development, market development, and diversification.

Next, we introduced the concept of the PLC and the notion that a different marketing mix is required for each of its four major phases. This idea was then linked to the proposal that firms need a product "portfolio" with a range of products at different complementary phases

of their life cycle as suggested by the "Boston Box."

Underlying a successful growth strategy is product (goods and/or services) *differentiation*, and we followed Deschamps and Nayak's five basic strategies for achieving this – through product proliferation, value, design, innovation, and service – and concluded by proposing how to leverage new-product growth and sustain differentiation.

Bibliography

Ansoff, H.I. (1957) Strategies for diversification. *Harvard Business Review*. **35** (2), 113–124.

Ansoff, H.I. (1968) *Corporate Strategy*, Pelican, Harmondsworth.

Baker, M.J. (2007) *Marketing Strategy and Management*, 4th edn, Palgrave Macmillan, Basingstoke.

Baker, M.J. and Hart, S.J. (1987) *Marketing and Competitive Success*, Philip Allen, London.

Baker, M.J. and Hart, S.J. (2007) *Product Strategy and Management*, 2nd edn, Pearson Education Ltd, Harlow.

Deschamps, J.-P. and Ranganath Nayak, P. (1993) 'Lessons from Product Juggernauts' in competing through products. *Prism*, Second Quarter.

Doyle P. (1999) Product life-cycle management, in *IEBM Encyclopedia of Marketing* (ed. M.J. Baker), Thomson Learning, London.

Doyle, P. (2000) *Value-based Marketing*, John Wiley & Sons, Ltd, Chichester. (A revised Second Edition published in 2008 covers the same issues in a modified form).

Firth, C. (1980) New approaches to strategic marketing planning, Unpublished MBA dissertation, Bradford University.

Henderson, B. (1968) *Perspectives on Experience*, Boston Consulting Group, Boston.

Johnson, S.C. and Jones, C. (1957) How to organise for new products. *Harvard Business Review*, **35**, 49–62.

Lele, M.M. (1992) *Creating Strategic Leverage*, John Wiley & Sons, Inc., New York.

Levitt, T. (1960) Marketing myopia. *Harvard Business Review*. **38** (4), 45–56.

Levitt, T. (1965) EXPLOIT the product life-cycle. *Harvard Business Review*. **43**, 81–94.

Senge, P.M. (1990) *The Fifth Discipline: The Art and Practice of the Learning Organisation*, Doubleday/Currency, New York.

The Conference Board (1974) Why Products Fail.

managing mature products

Susan Hart

Introduction

In his seminal article "Marketing Myopia," Theodore Levitt proclaimed that "every declining industry was once a growth industry," emphasizing the commercial imperative for management to scan the horizon for technological, social, and competitive developments that might erode the dominant or growth position of products at a given point in time (*see* GROWTH STRATEGIES). Indeed, there are numerous examples that the onslaught of intensive competition in the growth stage of a product's life cycle (PLC) often results in the onset of premature maturity for organizations that launched the innovation in the market. A pervasive characteristic of this tendency is that although industry-wide sales may still increase, the innovator's sales can flatten, requiring an alternative strategy from those used in previous stages of the PLC. The example of Intel illustrates the point well. Although a leader in terms of microprocessor market share, there persists a pattern of steeply declining prices for each generation of new processors (Hutt and Davidson, 2005). This indicates that, despite a market that is still growing, there are clear symptoms of product maturity accompanied by an alteration in marketing mix elements. But is price reduction the optimum response? Here, we explore what causes the slowdown in growth, the characteristics of maturity, and the options available in the mature phase of product and market development.

What is the Nature of Maturity and What are Its Causes?

Classical life cycle theory postulates that the market growth rate decreases as more and more of the available customers buy the product. When most of the available customers have bought the product (or service), that is, at the point of market saturation, the rate of growth decreases. At this juncture, sales continue to the point where sales volume is equal to the replacement rate, plus any natural growth in market size due to increase in population. This is typically the longest single stage of the product or the service life cycle. It is at this stage that products and services often become the "cash cows" and one feature of this is that they become the focus of the product and the brand managers whose existence is predicated upon the continuance of the product. The success of the product and the brand managers in prolonging this stage of the life of a product or service is the very reason critics of the PLC have dismissed the concept as not having any practical value. That said, classical PLC diagrams trace the progress of an innovation in the *absence* of managerial intervention. Its strength as a conceptual tool is that by considering the stage of a product or service in its life cycle, managers can define the strategic options available to them and so design appropriate strategies and tactics. *See* PRODUCT LIFE CYCLE for more details.

Another feature of slowing growth when a product approaches maturity is the profit *margin* reduction. The notion of margin is important here as mature products usually comprise the core of a company's product portfolio and hence the "cash cow" referred to above. However, during the elongated stage of the life cycle, margins can become increasingly narrow, which has been attributed by Rogers as long ago as 1962, to the following factors (*see* DIFFUSION OF INNOVATION):

1. increasing number of competitive products, leading to overcapacity and intensive competition;
2. market leaders under growing pressure from smaller companies;
3. strong increase in RESEARCH & DEVELOPMENT to find better versions of the product;
4. cost economies used up;
5. decline in product distinctiveness (*see* COMPETITIVE ADVANTAGE);
6. apathy and disenchantment in the distribution chain due to the product's declining sales; and
7. changing market composition where the loyalty of those first to adopt begins to waver.

These seven features are interrelated and often the tactics taken by firms accelerate

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

the transit of product markets from maturity to decline. The reaction of suppliers in a market where sales growth is declining, and there is increasing realization of the limits of growth, is countered by attempts to find ways around these limits. Typically, once growth in a market has slowed, further expansion can only come from intense competition for market share, using tactics such as product proliferation, discounting, own-label manufacture, and so on in an attempt to consolidate and protect their share. Global examples of these tactics in mature product markets include soft drinks, confectionery, and automobiles, whereas in services, telecom-provider competition has resulted in a wide array of service types, pricing strategies, and bundling techniques.

Although a prime characteristic of this stage in a product or service's life cycle, market saturation is not simply a case of having reached all potential consumers of the product in question. Sales' growth rates slow down even during the earlier stages of the PLC, before all potential users have tried the product. This is because, by the time laggards (the final tranche of customers to buy the product or the service) adopt the product or the service, those customers who bought the product early in its life cycle (termed "*innovators*" and "*early adopters*") are likely to have moved on, switching to other, possibly completely new, products. Thus, as the uptake of the product or the service decelerates, players in the distribution channel become less inclined to offer it, not only because of slower sales rates but also because of reduced margins compared to newer products still experiencing growth. Particularly in the case of physical products, where retail space is at a premium, retailers work to stock the assortment of product lines that optimize the return from the space available. Inevitably, slow-moving products will have their shelf and display space availability cut in favor of those whose growth curves suggest a fast-moving rate of sales and consequent higher margins. A good example of this can be found in the mobile phone handset margin, where specialist distributors exhibit new handset models to boost the uptake of particular service providers. This incurs a knock-on effect in that comparative lack of availability causes customers to buy the more easily available alternatives. Thus, a reduction in shelf and display space for mature products is likely to accelerate their demise. A further consequence is that the retailer may use the fact of dwindling sales to encourage tactics such as price cuts to incentivise purchase, reducing the profitability of the product even more. Put together, these tendencies can bring ever-intensifying pressure on mature brands.

In spite of these pressures, mature products are also known as *cash cows*, as for at least part of their maturity, they are considered to be assets to the firm and a necessary component of a balanced portfolio and have commanded much attention in terms of how best these assets can be managed. Indeed, most books on marketing management take the management of products and services in maturity as the "norm," discussing and exemplifying the permutations of the marketing mix and how they might be combined to best advantage. In this context, it is the development and marketing of new products, at best a solitary chapter in these books, that is singled out as a case differentiated from this norm. Such concentration has served to prolong the mature stage of the PLC beyond what could be regarded as its "natural" course, using two groups of strategies: offensive and defensive. Each is dealt with in turn below.

Strategies for Managing Maturity

Offensive strategies. An offensive or "take-off" strategy is, as the name implies, one designed to propel new energy into the sales of a mature product. The goal of such strategies is, in effect, to attempt to stimulate new growth in sales by new strategies. Various frameworks exist to categorize what these might entail. Two similar categorizations appeared within a few months of each other in the *Harvard Business Review* (September–October 1957 and May–June 1957). Ansoff's growth-vector matrix has become very well known – outlining the four *corporate* strategies of product penetration, product development, market development, and diversification (*see* PRODUCT DIVERSIFICATION). Johnston and Johnston's matrix, although less well known, is a 3 × 3 matrix, the unit of analyses of which are markets and technologies (both product and

process) offers a total of nine options as shown in Table 1. The original examples have been replaced by more contemporary instances from consumer markets and serve well as an organizing framework (*see also* GE/MCKINSEY MATRIX).

Essentially, this matrix uses three conditions of technology deployment (no technological change, improved technology, and new technology) and three market targeting conditions (no market change, strengthened market, and new market). Setting aside, for the purposes of this, the three options that call for a new technology (replacement, product-line extension, and diversification), which imply

Table 1 Johnston and Johnston matrix.

Increasing Technological Newness			
No Market Change		*Improved Technology*	*New Technology*
		Objective: to utilize more fully the company's present scientific knowledge and production	*Objective*: to acquire scientific knowledge and production skills
		Reformulation	Replacement
		To maintain an optimum balance of cost, quality, and availability in the formulas of present company products, for example, G 2-Screen Nintendo Lite	To seek new and better ingredients or formulation for presenting company products in technology not employed by the company now, for example, Playstation 3
Strengthened Market	*Remerchandizing*	*Improved Product*	*Product-line Extension*
Objective: to exploit more fully the existing markets for the company products	To increase sales to consumers of types now served by the company, for example, Coke; Pepsi, McDonald's, Dove	To improve present products for greater utility and products merchandisability to consumers, for example, Dove "Upside down" Deodorant	To broaden the line of products offered to present consumers through new technology, for example, Playstation portable (PSP); IPod Nano (video)
New Market	*New Use*	*Market Extension*	*Diversification*
Objective: to increase the number of types of consumers served by the company	To find new classes of consumers that can utilize present company products, for example, breakfast cereals as evening healthy snack; Nintendo for "brain train"	To reach new classes of consumers by modifying present products, for example, iPod shuffle to extend iPod use to younger age groups	To add to the classes of consumers served by developing new technical knowledge, for example, Nintendo Wii

Source: Adapted from Johnston and Jones (1957).

the development of another *new* product, leave us with five possible courses of action: remerchandizing, new use, reformulation, improved product, and market extension (*see* PRODUCT-LINE STRATEGIES). All of these may be considered as offensive strategies, although it is important to acknowledge that there may be instances of their deployment that also allow them to be identified with defensive strategies when used in combination with measures such as price adjustments or product augmentation (described later).

Remerchandizing. "Remerchandizing" can be achieved in a number of ways and has been long espoused by writers and practitioners alike. Drawing on some classical examples, Levitt's (1965) article on how to exploit the PLC used the example of the development of nylon by DuPont, to illustrate options such as promoting more frequent usage of the product among current users and developing more varied usage of the product among current users. At first glance, these options may seem identical, but they are nuanced. To explain further, it is interesting to turn to another classic piece, written by Twedt in 1964, which presents the concepts of heavy and light halves of usage concentration. Looking at the annual purchase concentrations in Table 2, Twedt categorized the annual sales among users and nonusers of the products. Using Lemon-Lime, the first item on the list, as an example, 42% of households are not consumers. Taking the households that do consume the product and dividing them into two groups, we see that the "light half" only accounts for 9% of consumption, whereas the "heavy half" accounts for a staggering 91%. This striking difference was found in many different categories, offering a choice to the mature product marketer – is resource to be spent on convincing heavy users to use still more of the product or on persuading the lighter users to increase their use. This reasoning illustrates the difference between the two remerchandizing options, namely, that more varied usage may be more appropriate for the heavy half, while more use may be more achievable with the light half, and each would require slightly different elements of the marketing mix. Returning to nylon, the choices facing marketers would have been the promotion of the product to light users (younger women, women in warmer climates) or to promote more varied use by heavier users by introducing product variants based on color and different textures. Consider the distinction between Diet Coke and Coke Zero, the former arguably merchandized for the heavy half (women) and zero for the light half (men). The factors influencing such a decision relate to the assessment of how difficult or problematic it is to increase usage among light users or nonusers. For example, Playboy's foray into purses, pencil cases, and other accessories for a young teen market, while commercially sensitive, has been much criticized as a way of remerchandizing the brand. Moreover, changing behavior, although central to the mission of marketing, is both complex and expensive, and the received wisdom is that reformulation of a product to new nonusers/new users' needs and wants is the better and more market-oriented course of action. Of course, alternative strategies are moving into new (geographical) markets that have previously not had access to the product – in other words, new market entry.

New uses. New uses of the existing product may be found, in view of the difficulties in persuading people to change behavior within the confines of current products available to them. The life cycle of nylon was extended by DuPont by modifying the basic material to a series of new uses – tire cord, textured yarns, carpet yarns, and so on. The illustration of this product in Levitt's article shows how restricting the original formulation to miscellaneous and circular knit applications, its sales by weight would have flattened out at 50 million pounds in 1962 as opposed to 10 times that amount across the totality of formulations of the material.

Other prominent examples of new uses include the application of Teflon (another DuPont product) beyond its original applications to include razor blades and the development of many uses of hemp from a fiber application (clothing and construction materials) to embrace specialty paper pulp, geotextiles, and cosmetic products.

Reformulation. Reformulation may be required to assist in the adoption of the new use, as well as to encourage current segments or markets to buy more products. The generational

Table 2 Annual purchase concentration in 18 product categories.

	Nonusers	*Users*	
Households		Light half	Heavy half
	42%	29%	29%
Lemon-lime	0	9%	91%
	22	39	39
Colas	0	10	90
	28	36	36
Concentrated frozen orange juice			
	0	11	89
	59	20	21
Bourbon	0	11	89
	54	23	23
Hair fixatives	0	12	88
	67	16	17
Beer	0	12	88
	67	16	17
Dog food	0	13	87
	52	24	24
Hair tonic	0	13	87
	4	48	48
Ready-to-eat cereals	0	13	87
	68	16	16
Canned hash	0	14	86
	27	36	37
Cake mixes	0	15	85
	3	48	49
Sausages	0	16	84
	11	44	45
Margarine	0	17	83
	34	33	33
Paper towels	0	17	83
	6	47	47
Bacon	0	18	82
	18	41	41
Shampoo	0	19	81
	2	49	49
Soaps and detergents	0	19	81
	2	49	49
Toilet tissue	0	26	74

Source: Twedt (1964).

development of mobile phone handsets as well as portable game consoles provides examples of the ongoing exploitation of a company's technological and market knowledge. In its most recent incarnation, the DS Lite harnesses multiple technologies such as wi-fi connectivity and touch-screen technology in a "double screen" variation on the original product.

Similarly, the inexorable changes in MP3 and MP4 are intended to position products across a variety of mobile-music consumer groups, including keep fit segments, commuters, and

home listeners. Design and styling can be used to greatly extend market use and even applications without requiring a fundamental change in the basic product. In addition to the changes in specification across the iPod range, color and styling have been particularly relevant and echo an article in *The Sunday Times* (18 December 1994), where Hugh Pearman exemplifies the concept and practice of reformulation through the changes made, over time, to the electric toaster. In his words,

> Toasters have been around as long as most people have had a domestic electricity supply, which is 80 years. The present generation is lined up on display in an electrical or department store near you. What does the act of choosing one say about our design awareness?
>
> All toasters are not exactly the same under the skin but they are as near the same that it makes no difference. They are boxes that nearly grill the bread, waffles, Pop-Tarts, or whatever between little electric fires and eject them just before they start to burn: an easy, well-proven technology, whether it is purely mechanical or microchip controlled. The last fundamental innovation in toaster design was in 1927, when the Sunbeam Company of America marketed the first pop-up model. Since then, there has been little to do, design-wise, except to alter the styling according to the tastes of the times. That meant round and bulbous in the 1950s, square and angular in the 1960s, decorated in the 1970s, elongated in the 1980s, and bulbous again in the 1990s.
>
> Designers try to give toasters the equivalent of sunroof and antilock brakes – wider slots, double slots, "cool wall" designs, and the like – but cannot get away from the fact that one needs only two controls: a push-down lever and a timer. Upgrades merely dress up a timeless concept and are anyway almost all adopted immediately by other manufacturers.'

The aim of the article is to illustrate that while the product remains functionally the same, it can be redesigned, styled, and reconfigured to encourage increased sales and sales to new and different market segments.

Product improvement. Comprehension of the PLC leads to the ability to preplan courses of action, which anticipate changes in competitive and market behavior and which when left unchecked could further slow down the growth of mature products. In consumer electronics, all the major players make use of product improvements, which are more substantial than the reformulations described above. Even in this high-technology product market, major technological breakthroughs can be easily copied, despite the use of patents. Patents themselves will often tell competitors all that they need to know to make a "me too" product. If a new product is not patented, then value analysis and competitive benchmarking of the new product soon reveal the majority of its secrets to others. It has been estimated that any major technological innovation is fully understood worldwide within 18 months of its first market entry. This can be readily seen in portable music products – in the 1980s, the Walkman was quickly replicated and in more recent times, MP3 clones have been introduced in quick succession.

To respond to these competitive pressures, key players in the market have used improvement strategies very effectively, by means of the introduction of frequent new products derived directly from the original product concept. For example, Sony introduced around 150 versions over a period of 10 years during the 1980s, ensuring that every possible niche and microsegment was covered, making it less economically viable for competitors to do the same. In a similar pattern, Sony introduces a major new feature on its camcorders at approximately six-month intervals. In both cases, the basic architecture of its products has remained constant, with each new product change launched in succession to keep a step ahead of key rivals. It must be noted that this strategy has its own limitations – it is about the improvement of a product category or technology where an organization can reap profits from fast, efficient, and knowledgeable exploitation of technologies in which it has developed considerable expertise. There is a risk, however, that concentrating on current technologies leads to blindness toward new technologies. It could be argued that Sony has not enjoyed continued level of success in the mobile-music market since the advent of MP3 technology, perhaps due

to a focus on old technology. Similarly, MP3/4 manufacturers will need to pay heed to the latest trends in the industry with an increasing number of wi-fi-enabled cellular phone handsets, which feature integrated music download features, set to double in volume by the end of 2010, compared to January 2008 (http://www.itfacts.biz). What must be guarded against, while keeping the product improvements going, is a myopic perspective on the next generation of music-listening techniques and devices.

Market extension. Market extension is a way of increasing sales of a product in a new market or market segment without a major development of the product itself. In reality, from a marketing perspective, some modification may be required to ensure that the marketing effort is not wasted by promoting a "production" orientation. Examples include the migration of breakfast cereals "positioning" from breakfast to snack foods bolstered by "weekly plans" for eating the product with different combinations as suggestions for times of the day other than breakfast, and the promotion of "brain training programs" for handheld gaming devices such as Nintendo, to extend their use from a youth/young adult market to more mature market segments. The toy and games market provides numerous examples of how traditional games, such as Monopoly, have extended markets for their products – for example, the junior market, as well as multiple boards for international appeal.

Thus far, the strategies discussed have been aimed at attracting *additional* customers to mature products. These are known as *offensive strategies* because they aim at taking sales from competitive products. It is also important, however, to continue to nurture current customers and mature–product market customers, so defensive strategies are also important; these are discussed in the following section.

Defensive strategies. During the later decades of the twenty-first century, levels of competition in mature markets became increasingly intense and with that dawned the realization that customer retention was equally – some contended more – important than customer creation. Where a market is saturated, by definition, all customers have purchased the product and so the only way to grow a customer base is by winning new customers at the expense of the competition. Not only is this a zero-sum game but it could also be portrayed as a negative-sum game because it costs more to create or win a customer than to keep one. The rule of thumb estimates confirm this in a number of ways:

- the need to spend 5 times as much on marketing to win a customer than that required to keep one happy;
- a 5% retention of the customer base that is normally lost through "churn" can improve profitability between 25 and 80%; and
- most customers will not switch from their current supplier for less than a 10% improvement in the value.

These alternative suggestions reinforce the adage that if marketing is "selling goods that do not come back to people who do," a large part of marketing effort must be about making sure those customers do, in fact, come back. In other words, it is repeat purchase and maintenance of the relationship, which is the key to long-run competitive advantage. It was this definition that F. Stewart De Bruicker and Gregory L. Summe probably had in mind when they wrote their article entitled "Make sure your customers keep coming back" in the Harvard Business Review article of 1985. De Bruicker and Summe address the question of how the onset of maturity affects buyer–seller relationships. They examine what steps sellers should take as buyers change from what they call "inexperienced specialists" to "experienced specialists." As customers gain experience with a product or service, their knowledge grows and attitudes can change, and the authors propose four strategies for dealing with this: strengthening account management practices, augmenting the product, improving customer services, and lowering prices.

While a technology associated with a new product or indeed a new product category is not well understood with the customer base, buyers are dependent upon sellers for advice on how to use a product and for after-sales service and support to get the maximum value from it. If customers become more familiar

with the product and it becomes more reliable, this dependency on the seller declines. As De Bruicker and Summe (1985) observe, "As customers gain familiarity with a product they find manufacturer's support programs to be of declining value. Their buying decisions become increasingly price-sensitive. They unbundle into components the products they once purchased as systems and open their doors to suppliers who sell on price and offer little in the way of product support. Even the most remote observer, once instructed can spot the pattern" (p. 93).

A firm that already enjoys the status of a supplier to a customer (also known as the *in-supplier*) usually has competitive advantage over the rival out-supplier, due to the perceived risks and costs of switching. Research amongst consumers suggests that they will often stick with a known product and supplier through habit and inertia even when they gain no particularly strong benefit or satisfaction from the habit. On the other hand, should something occur that causes active dissatisfaction, the habitual buyer may consider a change. Therefore, provided the in-supplier does not become complacent and continues to work at the relationship, using any of the four strategies cited earlier, it should be able to defend its position indefinitely.

Strengthening Account Management Practices. Strengthening account management practice directs resource and effort to maintain a multilevel relationship between seller and buyer and so keep purchasing decisions under review by general managers in senior positions. The involvement of top management is designed to prevent the product declining into commodity status when the purchaser (experienced specialist) becomes more knowledgeable and able to play one supplier off against another to secure the best price, delivery, and so on. In high-technology industries with rapid change and frequent new-product development, multilevel relationships are easier to develop and sustain. In fact, under the customer-active paradigm of new-product development, suppliers are frequently members of their customers' own product development team and so enjoy a privileged status. The challenge is to ensure the continuation of these relationships into the mature phase.

Product Augmentation. Product augmentation, the second defensive strategy, is described in some detail by Levitt (1980). He acknowledges that what constitutes product augmentation and added value may well vary from customer to customer, and the successful use of this approach will depend heavily on the seller's ability to identify courses of action that will enable the maintenance of differential competitive advantage and so avoid the rigors of price competition. There are numerous bases for augmentation – packaging changes, product delivery changes, changes in the distribution networks, and availability of the product to the customer. In recent years in the food industry, for example, repackaged products have stabilized falling sales in dairy, vegetable, and soft drinks. Hutt and Davidson (2005) explain how both Intel and Hewlett Packard have stayed competitive by focusing marketing on the additional value customers get from their technologies, rather than just the core attributes of the functionality of the products. A recent study reported in the *Sloan Management Review* described an example of augmentation in a Portuguese shoe manufacturing company that, by making business units more modular, they were able to respond to market niches more quickly and become a supplier of sample shoes to retail chains.

Customer Service Strategy. Customer service strategy, the third approach suggested by De Bruicker and Summe, is particularly appropriate for laggards who enter the market when the product is already mature. Although the product may have been around for quite some time, it is still new to the first-time buyer who may well have delayed entry because of "techno-fear". For such laggards, distribution through nonspecialist outlets that cater to novice buyers, customer training and education in how to use the product, and after-market support like IBM's Helpline will all add value for the new buyer.

Lowering Prices. Lowering prices is the response that may have to be offered when the product has become so standardized and commonplace that financial incentives or discounts will be essential to maintain buyer loyalty. Perversely, however, a very low price

may discourage buyers by increasing the buyer's perceived risk – "if it is that cheap it cannot be any good." Such an attitude is often found in the case of suppliers who offer much lower prices than their competitors. However, price competition can only be afforded by the few or be used as a short-term tactical weapon, as it erodes margins and ultimately destroys the cash income required for investment in future developments. The objective is to avoid price competition through a differentiation strategy: differentiation of anything – but especially the product.

Customer Relationship Management (CRM)

Customer relationship management (CRM) is a topic worth cross referencing in the context of managing mature products because, theoretically at least, it is an approach for closely monitoring customer buying behavior, developing and nurturing relationships with customers, and ensuring that the products and services on offer do match with customer's requirements. However, many commercial exchanges do not involve any particular relationship between seller and buyer other than a straightforward commercial transaction. Moreover, if a firm sets out to *manage* a relationship, this implies a desire to exercise power and control, which is antithetical to the very notion of a "relationship". These and other criticisms of CRM are beyond the scope of this article, but mention has been made because CRM is indeed a strategy suited to the maintenance of relationships, in the way that account management, described above, seeks to aid customer retention in mature markets.

Conclusion

As the market enters late maturity, the choice is seen as resting between withdrawal and growth. At this stage, the market is still very substantial and subject to little change. It offers little or no attraction for new entrants except for those with a radically new approach who may be prepared to attack the inertia and complacency of markets at this stage of their life cycle. Such attacks rarely come from within the industry. Business history is replete with examples of new technologies that displaced their predecessors: vinyl flooring displaced linoleum; cars and aircraft displaced trains and passenger ships; and MP3 technology displaced tape and disc technology. In each of these cases, the new product was not developed by those producing the incumbent products. In several cases, the opportunity did exist for the new product to be developed by those already enjoying market maturity but they either ignored the opportunity or did not wish to cannibalize their brands. Few reclaimed their prior position once they had allowed the new players to enter the market with their new products that satisfied demand. That is the picture for the entrenched supplier in a market that is in late maturity. Their job is twofold – manage maturity, keep the income flowing into the organization, but keep an eye on the future, monitor competitive threats, and be prepared to switch to new products and technologies once an irreversible decline sets in.

Bibliography

Ansoff, H.I. (1957) Strategies for diversification. *Harvard Business Review*, **35** (5), 113–124.

de Bruicker F.S. and Summe G.L. (1985) Make sure your customers keep coming back. *Harvard Business Review*, **63** (1), 92–98.

Hutt, K. and Davidson, A. (2005) Strategies for managing mature products. *Strategy and Leadership*, **33** (4), 51–52.

Johnston, S.C. and Jones, C. (1957) How to organize for new products. *Harvard Business Review*, **35**, 49–62.

Levitt, T. (1960) Marketing myopia. *Harvard Business Review*, **38** (4), 45–56.

Levitt, T. (1965) Exploit the product life cycle. *Harvard Business Review*, **43**, 81–94.

Levitt, T. (1980) Marketing success through differentiation – of anything. *Harvard Business Review*, **58** (1), 83–91.

Rogers, E.M. (1962) *The Diffusion of Innovations*, 3rd edn 1983; 4th edn 1995, Free Press, New York.

Twedt, B. (1964) How important to marketing strategy is the 'heavy user'? *Journal of Marketing*, **28** (1), 71–82.

leadership roles in product development

Abbie Griffin

Successful innovation requires multiple capabilities (Figure 1). Invention occurs in the FRONT END OF INNOVATION (FFE), before there is a well-formed concept. It is where technical possibilities are created and matched to customer problems and market needs, requiring technical competence, creativity, and market insight. Gaining project acceptance requires developing a business plan for the proposed product or service and shepherding that plan through the firm's funding and staffing process. This task fundamentally is one of managing the politics of the firm. Then, getting the product to market is a matter of project execution and implementation through formal processes, such as the Stage-Gate® (*see* THE STAGE-GATE IDEA TO LAUNCH SYSTEM).

Because invention, project definition and acceptance, and process execution each require different competencies, economic theory suggests that the most efficient management of the innovation process would be through a division of labor (*see also* ORGANIZING FOR INNOVATION) (Schumpeter, 1934). Indeed, in a traditional new-product development (NPD) setting, these stages are driven by *inventor* technologists, *champions*, and *implementers*, respectively, each with different sets of capabilities as identified in Table 1 (Griffin *et al*., 2007).

Definitions

Inventors typically have technical degrees, and frequently have advanced training and degrees in their field. These individuals have enough depth and breadth of technical understanding to solve critical problems and they exhibit creativity in the way they solve those technical problems (Tidd *et al*., 1997). They merit and seek a measure of autonomy in defining how they approach inventing (Kerr *et al*., 1977), have a commitment to their work and their profession, and may identify more with their profession and technical peers than with their firm. Inventors aspire to be ethical, unemotional, and unselfish in the conduct of their activities. Most frequently, inventors reside in the R&D laboratory, or in an engineering development group. While highly creative technically, inventors have little market or business knowledge, and frequently operate in isolation of others (O'Connor and Rice, 2005).

A *champion* puts himself on the line for an idea and uses any and every means of informal tactics and pressure to make it succeed (Schon, 1963). Champions usually do not create innovative ideas, but find them elsewhere in the firm (Markham and Aiman-Smith, 2001). Effective champions have technical competence, knowledge about the company, knowledge about the market, drive and aggressiveness, and political astuteness (Chakrabarti, 1974). They are most likely to use a transformational leadership style (Howell and Higgins, 1990).

Project managers, or *implementers*, organize the execution of the project and ensure that each task and milestone is completed on time, and within budget. They are task oriented (Turner and Miller, 2005), yet also have the ability to influence others on the team (Crawford, 2003). The most frequently used influence tactics by project managers are rational reasoning, consultation, and inspirational appeals (Lee and Sweeney, 2001). Higher project success is associated with project managers who exhibit higher proficiency in planning, a more participatory and motivating management style, and skills in the technology domain (Thieme *et al*., 2003).

While the vast majority of innovation projects divide the total labor among these different "specialist" individuals, some innovation projects are substantially the responsibility of just one individual, who has capabilities that allow him or her to drive all the phases of the project. These individuals, termed "*innovators*," invent, champion, and facilitate implementation. Innovators tend to be very driven systems thinkers, who are creative and curious about many different topics. They have a somewhat idealistic worldview – they want to solve important real-world problems that improve business outcomes and peoples' lives, where "important" is defined by people external to the corporation, like potential customers (Hebda *et al*., 2007). They take the perspective that technology is a means to an end, and that new technology must make money for the corporation. Innovators understand

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

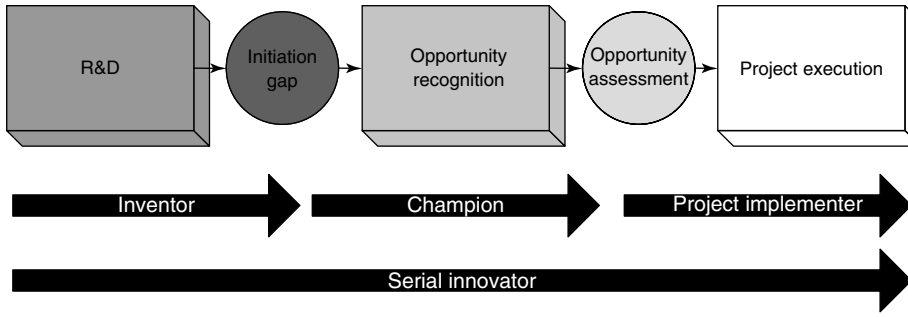


Figure 1 Specialist roles in the NPD process (reproduced from Griffin *et al*., (2007). © Wiley-Blackwell, 2007).

Table 1 New-product development roles and skills.

Core Skills	*Inventor*	*Champion*	*Implementer*	*Innovator*
Technical expertise	☑	●	●	☑
Market expertise		●		☑
Political guiding		☑		☑
Process implementation			☑	●

Reproduced from Griffin *et al*., (2007). © Wiley-Blackwell, 2007.
☑ primary skill; ● secondary sensitivity.

the reality of the political process of project acceptance. They prefer to use facts and other positive influencing mechanisms in managing the politics of innovation. Their process focuses on using extensive customer interaction to fully understand potentially interesting problems and extensive exploration of possible solutions, prior to launching into the implementation phase. They act in whatever role is needed at any particular point in time, much like entrepreneurs. Finally, innovators follow the new product into the marketplace, helping market and sell it. Overall, their process focuses on the tasks that are completed prior to a typical formal NPD implementation process, and extends beyond that process as well (Griffin *et al*., 2007).

Commentary

Table 2 summarizes the major differences among these types of NPD leaders. Because of their differences and the different specialist role that each prefers to take in innovation, it would seem that different management strategies need to be developed to allow each type of specialist to function in top form. It is unlikely that someone adept at managing inventors will be adept at managing champions, or that a good manager of implementers could also manage innovators. And yet, inventors, champions, and implementers need to interact effectively for a technology to proceed from capability to successful product in the marketplace.

Bibliography

Chakrabarti, A.K. (1974) The role of champion in product innovation. *California Management Review*, **17** (2, Winter), 58–62.

Crawford, L. H. (2003) Assessing and developing the project management competence of individuals, in *People in Project Management* (ed. J.R.Turner), Gower Press, Aldershot.

Griffin, A., Sim, E.W., Price, R.L., and Vojak, B.A. (2007) Exploring differences between inventors, champions, implementers and serial innovators in developing new products in large, mature firms. *Creativity and Innovation Management*, **16** (4), 422–436.

Hebda, J.M., Vojak, B.A., Griffin, A.J., and Price, R.L. (2007) The motivation of technical visionaries in large

Table 2 Personal and organizational differences across the roles.

	Inventor	*Champion*	*Implementer*	*Innovator*
Personality				
Creative ability	Technology	Business and political	Execution and coordination	Multi faceted
Openness	Strong introvert	Strong extrovert	Extrovert	–
Work style preference	Individual	Small group	Small group	Either
Orientation	Task	Communication	Task and people	Multiple
Perspective				
Childhood background	Well-educated family, early exposure to many topics	Lower-middle class	First-generation college	Traumatic childhood experiences
Worldview	Life long learning	Profit, success	Work to Live	Make the world a better place
Source of identity	Profession	Product	Firm, family	Customers
Knowledge Base from Preparation	Technology	Some technology, customer, and business	People and process	Technology, customer, market and business
Motivation	Technology is the end goal	Financial success and reputation	Earn a good living	Create solutions to customer problems
Attitude Toward Organizational Politics	Avoidance	Embrasure	Find rational answers	Acceptance – positive influencing
Process	Invention only	Political acceptance and high-level directing	Facilitate development after concept	Participate in all facets

Reproduced from Griffin *et al*., (2007). © Wiley-Blackwell, 2007.

American companies. *IEEE Transactions on Engineering Management*, **54** (3), 433–444.

Howell, J.M. and Higgins, C.A. (1990) Champions of technological innovation. *Administrative Science Quarterly*, **35** (2), 317–341.

Kerr, S., Von Glinow, M.A., and Schriesheim, J. (1977) Issues in the study of 'Professionals' in organizations: the case of scientists and engineers. *Organizational Behavior and Human Performance*, **18**, 329–345.

Lee, D.R. and Sweeney, P.J. (2001) An assessment of influence tactics used by project managers. *Engineering Management Journal*, **13** (2), 16–25.

Markham, S.K. and Aiman-Smith, L. (2001) Product champions: truths, myths and management. *Research Technology Management*, **44** (3), 44–50.

O'Connor, G.C. and Rice, M.P. (2005) Towards a Theory of New Market Creation for Radical Innovation, RPI Working Paper.

Schon, D.A. (1963) Champions for radical new inventions. *Harvard Business Review*, **41** (2), 77–86.

Schumpeter, J.A. (1934) *The Theory of Economic Development*, Harvard University Press, Cambridge.

Thieme, R.J., Michael Song, X., and Shin, G.-C. (2003) Project management characteristics and new product survival. *Journal of Product Innovation Management*, **20** (2), 104–119.

Tidd, J., Bessant, J., and Pavitt, K. (1997) *Managing Innovation*, John Wiley & Sons, Ltd, Chichester.

Turner, J.R. and Miller, R. (2005) The project manager's leadership style as a success factor on projects: a

literature review. *Project Management Journal*, **36** (2) 49–63.

Vojak, B., Griffin, A., Price, R.L., and Perlov, K. (2006) Characteristics of technical visionaries as perceived by American and British industrial physicists. *R&D Management*, **36** (1), 17–24.

new-product forecasting

Kenneth B. Kahn

New-product forecasting is the endeavor to come up with a meaningful estimate on which to guide new-product development and product launch decision making (*see* LAUNCH STRATEGIES). New-product forecasting is not necessarily a straightforward endeavor in practice and is often characterized as quite difficult (Mahajan and Wind, 1988) and error prone (Kahn, 2002). Employing appropriate techniques coupled with a well-structured new-product forecasting process appears to result in a greater propensity for new-product forecasting success. This emphasizes the need to recognize that new-product forecasting goes beyond technique application into process management (Kahn, 2006).

Types of New-Product Forecasts

Several types of new-product forecasts are possible and can be broken down in terms of *potential* versus *forecast*, and *market* versus *sales*. *Potential* represents a maximum attainable estimate, whereas *forecast* represents a likely attainable estimate. *Market* represents all companies within a given industry marketplace, whereas *sales* pertains to only the respective focal company. New-product forecast definitions are provided below:

- *Market potential*: the maximum estimate of total market volume reasonably attainable under a given set of conditions.
- *Sales potential*: the maximum estimate of company sales reasonably attainable within a given market under a given set of conditions.
- *Market forecast*: a reasonable estimate of market volume attainable by firms in that market under a given set of conditions.
- *Sales forecast*: a reasonable estimate of company sales attainable within a given market under a given set of conditions.

During the new-product forecasting effort, one or all of the above may be of interest. For example, early in the process, the forecasting focus is the market potential. Such forecasts are normally in dollars and are used to address the question of whether this is a good opportunity to pursue. During concept generation and pretechnical evaluation stages, forecasts investigate sales potential in answering the question of whether this is a good idea for the company to pursue. Entering technical development and launch phases, unit sales forecasts would become critical in order to plan for the launch and to ensure adequate supply through the channel. Product testing during technical development and market testing during the commercialization would help qualify key assumptions and better estimate unit demand and sales revenue.

Forecasting Techniques

There are numerous forecasting techniques available, with no one technique found to be the best. Among the multiple ways in which to categorize techniques, one way is to categorize forecasting techniques as judgmental techniques, quantitative techniques, and customer/market research techniques. Albeit showing only a sample of techniques, the below lists the more popular techniques associated with each of these three categories.

Judgmental techniques. Judgmental techniques represent techniques that attempt to turn experience, judgments, and intuition into formal forecasts. Six popular techniques within this category include jury of executive opinion, sales force composite, scenario analysis, Delphi method, decision trees, and assumptions-based modeling:

- *Jury of executive opinion*: a top-down forecasting technique where the forecast is arrived at through the ad hoc combination of opinions and predictions made by informed executives and experts.
- *Sales force composite*: a bottom-up forecasting technique where individuals (typically salespeople) provide their forecasts. These forecasts are then aggregated to calculate a higher level forecast.
- *Scenario analysis*: an analysis involving the development of scenarios to predict the future (Huss and Honton, 1987; Huss, 1988). Two types of scenario

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

analysis include exploratory and normative approaches. Exploratory scenario analysis starts in the present and moves out to the future based on the current trends. Normative scenario analysis leaps out to the future and works back to determine what should be done to achieve what is expected to occur.

- *Delphi method*: a technique based on subjective expert opinion gathered through several structured anonymous rounds of data collection (Woudenberg, 1991; Gordon, 1994). Each successive round provides consolidated feedback to the respondents, and the forecast is further refined. The objective of the Delphi method is to capture the advantages of multiple experts in a committee, while minimizing the effects of social pressure to agree with the majority, ego pressure to stick with one's original forecast despite new information, the influence of a repetitive argument, and the influence of a dominant individual.
- *Decision trees*: a probabilistic approach to forecasting where various contingencies and their associated probability of occurring are determined – typically in a subjective fashion. Conditional probabilities are then calculated, and the most probable events are identified (Ulvila, 1985).
- *Assumption-based modeling*: a technique that attempts to model the behavior of the relevant market environment by breaking the market down into market drivers. Then by assuming values for these drivers, forecasts are generated. These models are also referred to as *chain models* or *market-breakdown models* (Crawford and Di Benedetto, 2003; Kahn, 2007).

Quantitative techniques. Quantitative techniques are broken into the three subcategories of time series, "causal"/regression modeling, and other quantitative techniques. Time series techniques analyze sales data to detect historical "sales" patterns and construct a representative graph or formula to project sales into the future. Time series techniques used in association with new-product forecasting include

- *Trend line analysis*: a line is fit to a set of data. This is done either graphically or mathematically.
- *Moving average*: a technique that averages only a specified number of previous sales periods.
- *Exponential smoothing techniques*: a set of techniques that develop forecasts by addressing the forecast components of level, trend, seasonality, and cycle. Weights or smoothing coefficients for each of these components are determined statistically and are applied to "smooth" previous period information (Makridakis *et al.*, 1997; Mentzer and Bienstock, 1998).
- *Looks-like analysis (analogous forecasting)*: a technique that attempts to map sales of other products onto the product being forecast. Looks-like analysis is a popular technique applied to line extensions by using sales of previous product line introductions to profile sales of the new product.
- *Diffusion models*: models that estimate the growth rate of product sales by considering various factors influencing the consumer adoption process (*see* BASS MODEL). Considerations taken into account include the rate at which mass media (the coefficient of innovation) and word of mouth (the coefficient of imitation) affect lead users, early adopter, early majority, late majority, and laggard customer segments. Different types of diffusion models exist including the Bass model, the Gompertz curve, and the logistic curve, which are examples of some popular diffusion models. Diffusion models are also referred to as *technology S-curves* (*see* TECHNOLOGY S-CURVE) (Mahajan, Muller, and Bass, 1990; Lilien, Rangaswamy, and Van den Bulte, 1999; Mahajan, Muller, and Wind, 2000).
- *Autoregressive moving average (ARMA)/ Autoregressive integrated moving average (ARIMA) models*: a set of advanced statistical approaches to forecasting, which incorporate key elements of both time series and regression model building. Three basic activities (or stages) are considered: (i) identifying the model, (ii) determining the model's parameters, and (iii) testing/applying the model. Critical

in using any of these techniques is understanding the concepts of autocorrelation and differencing. ARMA/ARIMA models are also referred to as *Box–Jenkins techniques* (Makridakis *et al.*, 1997).

"Causal"/regression modeling techniques use exogenous or independent variables and through statistical methods, develop formulas correlating these with a dependent variable. The term "causal" is very loosely used because these models are predicated on correlational relationships and not true cause and effect relationships. Four popular techniques within this subcategory include

- *Linear regression*: a statistical methodology that assesses the relation between one or more managerial variables and a dependent variable (sales), strictly assuming that these relationships are linear in nature (Neter *et al.*, 1996). For example, price may be an important driver of new-product sales. The relationship between price and the quantity sold would be determined from prior data of other products within the product line and then used to predict sales for the forthcoming product.
- *Nonlinear regression*: a statistical methodology that assesses the relation between one or more managerial variables and a dependent variable (sales), but these relationships are *not* necessarily assumed to be linear in nature.
- *Logistic regression*: a statistical methodology that assesses the relation between one or more managerial variables and a binary outcome, such as purchase versus nonpurchase. A logistic regression model calculates the probability of an event occurring or not occurring.
- *Event modeling*: often a regression-based methodology that assesses the relation between one or more events, whether company initiated or nonaffiliated with the company, and a dependent variable (sales). For example, a promotion used with prior product launches would be analyzed and the bump in sales caused by this promotion statistically determined. The expected bump in sales would be correspondingly mapped to the sales of the new product.

The other category contains those techniques that employ unique methodologies or represent a hybrid of time series and regression techniques. A sample of these forecasting techniques include

- *Expert systems*: typically computer-based heuristics or rules for forecasting. These rules are determined by interviewing forecasting experts and then constructing "if-then" statements. Forecasts are generated by going through various applicable "if-then" statements until all statements have been considered.
- *Neural networks*: advanced statistical models that attempt to decipher patterns in a particular sales time series (Adya and Collopy, 1998). These models can be time consuming to build, difficult to explain, and in many cases, the models are proprietary.
- *Simulation*: an approach to incorporate market forces into a decision model. "What if" scenarios are then considered. Normally, simulation is computer based. A typical simulation model is Monte Carlo simulation, which employs randomly generated events to drive the model and assess outcomes.

Customer/market research techniques. Customer/market research techniques include those approaches, which collect data on the customer/market and then systematically analyze these data to draw inferences on which to make the forecasts. Four general classes of customer/market research techniques include

- *Concept testing*: a process by which customers (current and/or potential customers) evaluate a new-product concept and give their opinions on whether the concept is something that they might have interest in and would likely buy (*see* CONCEPT TESTING). The purpose of concept testing is to proof the new-product concept.
- *Product use testing*: a process by which customers (current and/or potential customers) evaluate a product's functional characteristics and performance (*see* PRODUCT TESTING). The purpose of product use testing is to proof the product's function.

- *Market testing*: a process by which targeted customers evaluate the marketing plan for a new product in a market setting. The purpose of market testing is to proof the proposed marketing plan and the "final" new product.
- *Premarket testing (*see PRETEST MARKET MODELS*)*: a procedure that uses syndicated data and primary consumer research to estimate the sales potential of new-product initiatives (Urban *et al.*, 1983; Fader and Hardie, 2001). Assessor and BASES are two proprietary new-product forecasting models associated with premarket testing. BASES is commonly employed in the consumer products goods industry.

While there are a number of forecasting techniques available, it is important to realize that not all are appropriate for every forecasting situation. Qualitative techniques are quite adaptable, but very time consuming; they would not be appropriate in situations where a severe time constraint exists. Quantitative techniques require data, and rely on the critical assumption that current data will correspond to future states; if these are not feasible, quantitative techniques would not be meaningful. Customer/market research tools are time consuming and expensive to perform. Budget constraints could seriously hamper what degree of customer/market research may be applied. A "toolbox" approach is therefore recommended for applying new-product forecasting techniques.

New-Product Forecasting Accuracy

Even with a plethora of techniques, new-product forecasting is characteristically associated with low accuracy (high forecast error). Research shows that the overall average accuracy across the six types of new products is around 58%, with the accuracy of cost improvements generally at 72%; product improvement forecasts at 65%; line-extension forecasts at 63%; market extension forecasts at 54%; new-category entry (new-to-the-company) forecasts at 47%; and new-to-the-world products at 40% (Kahn (2002); these mean values of new-product forecasting accuracy were collected by asking respondents to indicate the average forecast accuracy achieved one year postlaunch). The nature of these accuracies data reaffirms that newer markets are more troublesome to forecast (i.e., market extensions, new-category entries, and new-to-the-world products), than those situations where a current market is being served (i.e., cost improvements, product improvements, and line extensions).

One of the underlying reasons for low accuracy is the multiple organizational biases that impinge on the new-product forecasting endeavor and cause unnecessary error. Naturally, a new-product forecast should be developed with a keen eye toward realism and presented to management regardless of whatever outcome shown, but seldom does this happen. Various research studies have documented the impact of organizational biases on new-product forecasting (Tyebjee, 1987; Forlani *et al.*, 2002; Bolton, 2003). While it is very unlikely that new-product forecasting will be free of all biases owing to the need to rely on judgment, understanding the persistent types of biases can allow for necessary procedures that enable transparency of new-product forecasting figures and sounder new-product decision making. When "systematic biases" are identified and properly mitigated, the chance for a new-product forecast to be more "on target" increases.

The New-Product Forecasting Process

Applying techniques appropriately and addressing organizational biases that surround new-product forecasting mandate a process approach toward new-product forecasting. In doing so, the new-product forecasting process should highlight critical uncertainties inherent in developing the new-product forecast such as cannibalization effects and market penetration. The process also should designate what data need to be collected to address these uncertainties. A process approach to new-product forecasting further mandates the continuous building on experiences from prior new-product forecasts, cross-functional communication (especially with marketing), and customer feedback, thereby enabling

organizational learning and understanding on which to make a credible and realistic forecast.

Assumptions management is an important part of the new-product forecasting process. Accordingly, the new-product forecasting process should clearly specify assumptions and make them transparent so that there is company understanding of what underlies these assumptions. After launch, a tracking system would closely monitor and control these assumptions to determine if forecasts will come to fruition or deviate significantly.

As part of a successful new-product-forecasting endeavor, successful companies have realized that new-product forecasts should not be point forecasts, but rather, be presented as ranges. These ranges typically become more narrowed as the product approaches and enters the launch phase. For example, pessimistic, likely, and optimistic cases could be connected with the monitoring and control of assumptions to determine which scenario is playing out. Best-in-class companies also construct databases to collect, track, and reflect on new-product rollouts, especially those in consumer packaged goods industries. Such internal databases are crucial for the validation of new-product forecasting assumptions, as well as for documenting new-product forecasting accuracy. Tying the internal database with syndicated data, market share, and competitor data can enable more robust analyses related to brand preferences, price elasticities, and geographic rollout scenarios.

Bibliography

Adya, M. and Collopy, F. (1998) How effective are neural nets at forecasting and prediction? A review and evaluation. *Journal of Forecasting*, **17**, 451–461.

Bolton, L.E. (2003) Stickier priors: the effects of nonanalytic and analytic thinking in new product forecasting. *Journal of Marketing Research*, **XL**, 65–79.

Crawford, M. and Di Benedetto, A. (2003) *New Products Management*, McGraw-Hill/Irwin, Boston.

Fader, P.S. and Hardie, B.G.S. (2001) Forecasting trial sales of new consumer packaged goods, in *Principles of Forecasting: A Handbook for Researchers and Practitioners* (ed. J. Scott Armstrong), Kluwer Academic Publishers, Norwell, pp. 613–630.

Forlani, D., Mullins, J.W., and Walker, O.C. Jr (2002) New product decision making: how chance and size of loss influence what marketing managers see and do. *Psychology and Marketing*, **19** (11), 957–981.

Gordon, T.J. (1994) The Delphi Method, AC/UNU Millennium Project, Futures Research Methodology, Summary Report, 30 pp.

Huss, W.R. (1988) A move toward scenario analysis. *International Journal of Forecasting*, **4** (3), 377–388.

Huss, W.R. and Honton, E.J. (1987) Scenario planning—what style should you use? *Long Range Planning*, **20** (4), 21–29.

Kahn, K.B. (2002) An exploratory investigation of new product forecasting practices. *Journal of Product Innovation Management*, **19**, 133–143.

Kahn, K.B. (2006) *New Product Forecasting: An Applied Approach*, ME Sharpe, Armonk.

Kahn, K.B. (2007) Using assumptions-based models to forecast new product introduction, *The PDMA Toolbook for New Product Development*, John Wiley and Sons, Inc., Hoboken, pp. 257–272.

Lilien, G., Rangaswamy, A., and Van den Bulte, C. (1999) Diffusion Models: Managerial Applications and Software, Institute for the Study of Business Markets Report #7-1999.

Mahajan, V., Muller, E., and Bass, F.M. (1990) New product diffusion models in marketing: a review and directions for research. *Journal of Marketing*, **54** (1), 1–26.

Mahajan, V., Muller, E., and Wind, Y. (2000) *New-Product Diffusion Models*, International Series in Quantitative Marketing, vol. 11, Kluwer Academic Publishers, Boston.

Mahajan, V. and Wind, Y. (1988) New product forecasting models: directions for research and implementation. *International Journal of Forecasting*, **4** (3), 341–358.

Makridakis, S.G., Wheelwright, S.C., and Hyndman, R.J. (1997) *Forecasting: Methods and Applications*, John Wiley & Sons, Inc., New York.

Mentzer, J.T. and Bienstock, C.C. (1998) *Sales Forecasting Management*, Sage Publications, Thousand Oaks.

Neter, J., Kutner, M.H., Nachsheim, C.J., and Wasserman, W. (1996) *Applied Linear Regression Models*, Irwin, Chicago.

Tyebjee, T.T. (1987) Behavioral biases in new product forecasting. *International Journal of Forecasting*, **3**, 393–404.

Ulvila, J.W. (1985) Decision trees for forecasting. *Journal of Forecasting*, **4** (4), 377–385.

Urban, G.L., Katz, G.M., Hatch, T.E., and Silk, A.J. (1983) The ASSESSOR pre-test market evaluation system. *Interfaces*, **13** (6), 38–59.

Woudenberg, F. (1991) An evaluation of Delphi. *Technological Forecasting and Social Change*, **40** (2), 131–150.

innovation typologies

Rajesh K. Chandy and Jaideep C. Prabhu

Innovation, like love, is a many-splendored thing. Whether in the form of a shiny new gadget sought out by eager consumers, or a new technique that saves businesses time and money, or a new service that destroys the value of existing services, innovations come in many forms and affect our lives in different ways. This article examines the dimensions of innovation, and presents a means of classifying innovations.

By definition, all innovations have an element of novelty in them: they are different from what existed before. Yet, all innovations are not the same. Indeed, scholars have over the years proposed a number of different innovation typologies, or ways of classifying innovations. By distinguishing among the many typologies of innovations, we can better manage their development and better predict their sources and their impact. To this end, this article first describes and defines some of the prominent typologies of innovations that scholars have highlighted in the literature. These typologies have generally existed in isolation from each other in the literature. We note some of the commonalities among these typologies, and then propose a metatypology (or typology of typologies) that might serve as a unifying framework for the study and management of innovation (*see also* RADICAL INNOVATION; PRODUCT PLATFORMS; PRODUCT MODULARITY).

What Types of Innovation Exist in the Literature?

Table 1 presents a list of the prominent types of innovation that have been described in the literature. In this section, we define each of the types of innovation listed in Table 1, and illustrate them with examples. This section highlights the sheer variety of innovations that exist, and have been studied in the literature.

A product innovation involves the commercial introduction of a product that is new to customers (Schumpeter, 1934). A service innovation involves the commercial introduction of a service that is new to customers (*see* SERVICE INNOVATION MANAGEMENT). The introduction of the Kodak Brownie in 1900 is an example of a product innovation: it offered cheap and easy use of photography to adults and children alike. FedEx and UPS are examples of service innovations: they enabled customers to deliver documents and packages overnight. A process innovation is one that involves the use of a new approach to creating or commercializing products or services (*see* PROCESS INNOVATION). An example of a process innovation is Henry Ford's use of the assembly line in the manufacture of automobiles.

A technological breakthrough is a product, service, or process that involves scientific principles that are substantially different from those of existing products, services, or processes (Chandy and Tellis, 1998). An alternate label for such an innovation is platform innovation (Sood and Tellis, 2005). The first electronic cameras were technological breakthroughs, since they used a core technology – electronic imaging – to take photographs, in contrast to existing cameras that relied on celluloid roll technology to take photographs.

A component innovation is a new product, service, or process that uses new parts, modules, or materials, but relies on the same core technology as existing products, services, or processes (Tellis and Sood, 2008). For example, in the field of data recording, magnetic tapes and floppy disks differed in the components and materials they used, though they were both based on the core technology of magnetic recording.

An architectural innovation (Henderson and Clark, 1990) or design innovation (Tellis and Sood, 2008) involves the reconfiguration of the linkages and layout of components, but relies on the same core technology as existing products, services, or processes. For example, floppy disks decreased from 14 to 8 inches in 1978, to 5.25 inches in 1980, to 3.5 inches in 1985, and to 2.5 inches in 1989, though each was based on magnetic recording (Christensen, 1997; Tellis and Sood, 2008).

A business model innovation involves systemic changes to the value proposition offered by a product or service, and to the cost structure incurred by the firm offering it (Velu, Prabhu, and Chandy, 2009). These changes are systemic as they involve changing

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 A Meta-Typology of Innovation.

		Dimension of Novelty		
		Concept	*Customer*	*Company*
Feature	Attributes	Product/service/ process innovation	Market breakthrough	New to the firm innovation
		Technological break-through/platform innovation	Disruptive innovation	Organizational/ administrative/ management innovation
		Component innovation		
		Architectural/design innovation		
		Business model innovation		
	Effects	Drastic innovation	Discontinuous innovation	Competence-destroying innovation
		Revolutionary innovation	Disruptive innovation	Competence-enhancing innovation
				Disruptive innovation

multiple elements of the marketing mix and cost structure at the same time. An example of a business model innovation is Amazon.com. Amazon.com is an innovation that, relative to bricks and mortar bookstores, involves changes to the product (a far greater assortment than any bricks and mortar book retailer), distribution (books available everywhere all the time through the Internet), price (lower prices in general), and promotion (access to online customer reviews). It also involved a reduction in the cost of the firm's offering through lower overheads and greater economies of scale.

A drastic innovation (Reinganum, 1985) or revolutionary innovation (Caselli, 1999; Rosen, 1991) is one that makes current products obsolete. For example, electronic calculators made slide rules obsolete.

An innovation that is a market breakthrough provides substantially higher customer benefits per dollar than existing products, services, or processes, but is based on the same core technology as existing products, services, or processes (Chandy and Tellis, 1998). An example is cable television with signal compression technology that allows a significantly larger number of channels to be transmitted to customers than before, while using existing cable technology.

A radical innovation is one that employs substantially new technology and offers substantially higher customer or user benefits relative to existing products, services, or processes (Sorescu, Chandy, and Prabhu, 2003). A radical innovation combines the technological novelty of technology breakthroughs, and the dramatic improvements in customer and user benefits evident in market breakthroughs. As such, radical innovations tend to be fairly rare. However, they can have powerful effects on the fates of individual firms and entire industries (Chandy and Tellis, 1998, 2000). An incremental innovation is one that involves only minor changes to technology or minor improvements in benefits.

A disruptive innovation (Govindarajan and Kopalle, 2006a, p. 15; Christensen, 1997) introduces a different set of features, performance, and price attributes relative to existing products, making it an unattractive combination for mainstream customers at the time of product introduction because of inferior performance on the attributes these customers value a high price – although a different customer segment may value the new attributes. Subsequent developments over time, however, raise the new product's attributes to a level sufficient to

satisfy mainstream customers, thus attracting more of the mainstream market (Govindarajan and Kopalle, 2006a, p. 15). As Govindarajan and Kopalle (2006b, p. 199) note, "Canon's introduction of slower but inexpensive tabletop photocopiers in the late 1970s relative to Xerox's high-speed big copiers is an example of disruptive innovation. The tabletop copiers were rapidly accepted by small businesses and individuals who appreciated the convenience and price despite poor resolution. At the time of their introduction, the mainstream market (larger companies) still preferred the large copiers because of speed, resolution, collation, and so on. However, over time, further developments in small copiers have allowed Canon to improve quality, speed, and features and offer them at a price point that is sufficient to satisfy the needs of mainstream market."

A discontinuous innovation is one that requires customers to establish different behavior patterns (Robertson, 1967, p. 16). It alters existing patterns of use or creates new patterns of use (Robertson and Gatignon, 1986). Typewriters and microwaveable foods were discontinuous innovations, since they changed longstanding patterns of use and behavior among customers.

A new to the firm innovation refers to the adoption of a product, idea, or behavior by a firm that had not previously adopted the innovation (Booz, Allen, Hamilton, 1982). New to the firm innovations that involve the adoption of new management principles, processes, and practices are referred to as *administrative innovations* (Damanpour, 1991) or *organizational innovations* (Hage, 1999), or *management innovations* (Birkinshaw, Hamel, and Mol, 2008). For example, the adoption of electronic banking by financial institutions and of total quality management practices by manufacturing firms both represent new to the firm innovations. A new to the world innovation, in contrast, represents the first time an innovation is introduced commercially in any marketplace (Maclaurin, 1950).

A competence-destroying innovation is one that requires new skills, abilities, and knowledge in the development and production of a product relative to those held by existing firms in an industry (Tushman and Anderson, 1986, p. 442). As Tushman and Anderson (1986) note, a competence-destroying innovation either creates a new product class (e.g., xerography or automobiles) or substitutes for an existing product (e.g., diesel vs steam locomotives; transistors vs vacuum tubes).

A competence-enhancing innovation is an order of magnitude improvement in price, performance, or efficiency that builds on existing know-how within a product or process class (Tushman and Anderson, 1986, p. 442). Such innovations substitute for older technologies, yet do not render obsolete skills required to master the old technologies. For example, electric typewriters represented a competence-enhancing innovation relative to mechanical typewriters.

What Dimensions Underlie These Varied Types of Innovations?

The extensive literature on innovation typologies can sometimes resemble a Tower of Babel with different terms used to refer to similar concepts and, conversely, similar terms used to refer to different concepts. Moreover, terms and schemes sometimes overlap or only partially capture the complexity of the phenomenon of innovation. In this section, we present a metatypology, or typology of typologies, that seeks to provide some structure to the existing typologies in the literature. In doing so, we follow three criteria. First, the typology must be exhaustive and comprehensively capture the full extent of a phenomenon. Second, the typological units or categories of the typology should be mutually exclusive, thus avoiding redundancies and overlap. Finally, the typology should be parsimonious.

Taking into account these three hallmarks, we synthesize the various schemas discussed in the section above into a single typology laid out in Table 1. We identify two key dimensions that underpin all the various typologies mentioned in the literature. The first dimension classifies innovations based on attributes (i.e., what they are) versus their effect (i.e., what they do). The second dimension classifies innovations based on their locus of novelty: the novelty of the concept itself, or the novelty of the innovation in relation to the customer or the company that introduces or adopts it. Accordingly, we classify

product/service/process innovation, technological breakthrough/platform innovation, component innovation, architectural/design innovation, and business model innovation all as types of innovation that emphasize the innovation's attributes (i.e., what it is) and whose locus of novelty lies in the concept behind the innovation. Similarly, we classify market breakthroughs and disruptive innovations as types of innovation that emphasize the innovation's attributes (i.e., what it is) and the novelty of the innovation in relation to the customer. Likewise, we classify new to the firm innovation, organizational innovation, administrative innovation, and management innovation as types of innovation that emphasize the innovation's attributes (i.e., what it is) and the novelty of the innovation in relation to the company. Drastic innovation and revolutionary innovation we classify as being types of innovations that emphasize the innovation's effects (i.e., what it does) and whose locus of novelty lies in the concept behind the innovation. In contrast, we classify discontinuous innovation and disruptive innovation as types of innovation that emphasize the innovation's effects (i.e., what it does) and the novelty of the innovation in relation to the customer. And finally, we classify competence-destroying innovation, competence-enhancing innovation, and disruptive innovation as types of innovation that emphasize the innovation's effects (i.e., what it does) and the novelty of the innovation in relation to the company.

What are the Implications of the Typology We Propose?

The metatypology in Table 1 has several implications for research and practice. We discuss each of these in turn.

Implications for research. The metatypology we propose in Table 1 helps to reduce some of the confusion and complexity in the literature on the types of innovation. While the literature suggests many different dimensions underlie innovation, the typology shows that, in fact, only two major dimensions underlie most types of innovation. In identifying only two underlying dimensions, the typology meets the important condition of parsimony in a classification scheme. Further, by capturing most major types of innovation, the typology is also exhaustive and comprehensive. Third, the typology ensures that the categories it suggests are mutually exclusive, thus avoiding redundancies and overlap.

By meeting the major conditions of a good typology, Table 1 provides the basis for a rigorous approach to the study of innovations. Specifically, it forms the basis for further research that can help better circumscribe the rich and diverse phenomenon of innovation, and its antecedents and consequences. Most importantly, it furthers the task of good theory in the area, namely, to be able to explain and predict aspects of the phenomena in a parsimonious and logically rigorous way.

Implications for practice. The typology in Table 1 provides at least three major benefits for managers. Specifically, it helps managers make better innovation decisions in the areas of (i) portfolio assessment, (ii) resource allocation (time, people, money), and (iii) portfolio management. A key issue for managers is to be able to assess in real time the extent to which their portfolio of existing products as well as those in the pipeline are complete and sufficient given their firms' short and long term objectives. The typology in Table 1 enables managers to keep a track of the various types of innovations they have planned and introduced in markets over time. This in turn helps them assess any imbalances in the portfolio. Any imbalances in the portfolio would immediately suggest a reallocation of resources devoted to the development and support of innovations. Such imbalances would also suggest means to systematically identify and generate new innovations by, for instance, seeking novelty in innovations based on their relationship with concepts, customers, or the company itself.

Bibliography

Ashish, S. and Tellis, G.J. (2005) Technological Evolution and Radical Innovations. *Journal of Marketing*, **69** (3), 152–168.

Birkinshaw, J., Hamel, G., and Mol, M. (2008) Management innovation. *Academy of Management Review*, **33** (4), 825–845.

Booz, Allen, Hamilton (1982) *New Product Development in the 1980's* Booz, Allen and Hamilton, New York.

Caselli, F. (1999) Technological revolutions. *American Economic Review*, **89** (1), 78–102.

Chandy, R.K. and Tellis, G.J. (1998) Organizing for radical product innovation: the overlooked role of willingness to cannibalize. *Journal of Marketing Research*, 474–487.

Chandy, R.K. and Tellis, G.J. (2000) The incumbent's curse? Incumbency, size, and radical product innovation. *Journal of Marketing*, **64** (3), 1–17.

Charitou, C.D. and Markides, C.C. (2003) Responses to disruptive strategic innovation. *Sloan Management Review*, **44** (2), 55–63.

Christensen, C. (1997) *The Innovator's Dilemma*, Harvard Business School Press, Boston.

Christensen, C.M. (1997) *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*, Harvard Business School Press, Boston, MA.

Damanpour, F. (1991) *Organizational Innovation: A Meta-Analysis of Effects of Determinants and Moderators*, *Academy of Management Journal*, **34**, 555–590.

Govindarajan, V. and Kopalle, P.K. (2006a) The usefulness of measuring disruptive innovations ex-post in making ex ante predictions. *Journal of Product Innovation Management*, **23**, 12–18.

Govindarajan, V. and Kopalle, P.K. (2006b) Disruptiveness of innovations: measurement and an assessment of reliability and validity. *Strategic Management Journal*, **27** (2), 189–199.

Hage, J.T. (1999) Organizational innovation and organizational change. *Annual Review of Sociology*, **25**, 597–622.

Henderson, R. and Clark, K. (1990) Architectural innovation: the reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly*, **35** (1), 81–112.

Maclaurin, W.R. (1950) The process of technological innovation: the launching of a new scientific industry. *American Economic Review*, **40** (1), 90–112.

Reinganum, J.F. (1985) Innovation and industry evolution. *Quarterly Journal of Economics*, **100** (1), 81–99.

Robertson, T.S. (1967) The process of innovation and the diffusion of innovation. *Journal of Marketing*, **31** (1), 14–19.

Robertson, T.S. and Gatignon, H. (1986) Competitive effects on technology diffusion competitive effects on technology diffusion. *Journal of Marketing*, **50** (3), 1–12.

Rosen, R.J. (1991) Research and development with asymmetric firm sizes. *RAND Journal of Economics*, **22** (3), 411–429.

Schumpeter, J. (1934) *The Theory of Economic Development*, Harvard University Press, Boston.

Sood, A. and Tellis, G. (2005) Technological evolution and radical innovations. *Journal of Marketing*, **69** (3), 152–168.

Sorescu, A., Chandy, R., and Prabhu, J. (2003) Sources and financial consequences of radical innovation: insights from pharmaceuticals. *Journal of Marketing*, **67** (4), 82–102.

Tushman, M.L. and Anderson, P. (1986) Technological discontinuities and organizational environments. *Administrative Science Quarterly*, **31** (3), 439–465.

Velu, C., Prabhu, J., and Chandy, R. (2009) Business Model Innovation in Network Markets, working paper.

innovation metrics

David Reibstein and Venkatesh Shankar

Introduction

Innovation is essential for the future of any firm. Philips, for example, in 2008 had 58% of their sales from newly introduced products – products introduced within the last year (for Business to Consumer [B2C] products) or three years (for Business to Business [B2B] products) (Philips, 2009). This is an amazingly high number for a large established company such as Philips. Nonetheless, it is essential that the flow of new products be constant to keep the firm vibrant and reinvent itself. To achieve these objectives, firms make significant investments in innovation. To manage and track these investments, they gather numerous measures and monitor their progress. Many times, these metrics are used to diagnose the innovation investment performance. There are various stages of innovation development, and not surprisingly there are measures at each of the different stages (*see* THE STAGE-GATE IDEA TO LAUNCH SYSTEM). We will review the innovation measures that are regularly collected and used to assess the innovation process. Ideally, these measures should provide strategic direction for the innovation activities and guide future resource allocation decisions. They should also help in assessing the effectiveness of innovation spending, and, in many cases, help make managers accountable.

Stages of Innovation Development

Firms go through several stages of innovation development, starting from the inception of an idea to the launch of a product or service. Firms need to assess performance at each of these stages. To facilitate such a continual assessment, firms can benefit from the use of an "innovation dashboard" that offers a snapshot of performance at each major innovation development stage – from the inputs, through the development process, all the way to outputs such as introduction into the marketplace (see Table 1). In the section that follows, we will review the measures commonly used to assess progress at each stage.

Input metrics. These are measures that capture all early stage investments and resources required during the inception of the innovation. It is essential for a firm to assess whether there is a sufficient number of new product ideas (pipeline of ideas) and whether a sufficient amount is being budgeted for future potential. It would be possible to assess where the new product ideas are coming from and if there is a sufficient flow of innovations to ensure future viability for the firm. It is also necessary to keep track of the resources that are required at the early stage. There needs to be a steady flow of innovations at various stages of development (*see* PORTFOLIO MANAGEMENT). If the flow of innovations is out of balance, this will be reflected in the financial performance of the firm. If there is an imbalance in projects such that most investments are skewed toward early stage innovations, it could drain the resources of the firm, without letting it reap the rewards of their introductions to the market. Further, there may not be enough resources to bring other projects to completion. At the input stage, there tends to be measures of the number of projects under development and the corresponding resources they require.

Process effectiveness metrics. These measures relate to the effectiveness of the research and development process. It is natural for many projects to be initiated, but for some to be terminated at different points during the development process. As firms develop new innovation projects, they continually undertake market assessments for those projects. If a large percentage of the projects make it through to completion, then the idea generation process is probably not producing concepts that stretch the firm very much. Obviously, the other extreme is probably even more troublesome. It would mean that many resources are being spent on projects that never make it to the market. One of the key measures is the number of projects reaching the different stages (steps of development) on time. Firms typically follow a timetable of stages for each innovation project. If an insufficient number of innovations are meeting their scheduled gates on time, it would suggest that either the plans or the forecasts need to be modified. Alternatively, it could

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Innovation dashboard.

Inputs	*Process effectiveness*	*Performance outcomes*
1. R&D spending (percent of sales) 2. Human resources devoted to innovation – Number of personnel on innovation projects – Payroll costs of innovation employees 3. Number of new ideas/concepts in the pipeline 4. Number of R&D projects in active development 5. Percent of ideas/concepts from outside the firm	1. Development activities – Percentage hitting gates on time – Percentage meeting quality guidelines 2. Patenting activity – Number filed – Number commercialized – Percentage of new ideas covered by patents 3. Budget versus actual – Time – Cost/investment 4. Number of new products introduced (firm innovativeness) 5. Average time to market 6. Percentage of projects that are major improvements 7. Percentage of projects resulting in new product launch	1. Percentage of sales from new products in past *N* years 2. Successful new products (number and ratio of new products meeting financial goals) 3. Revenue growth 4. Return on investment in innovation 5. Percentage of profits from new customers (or occasions) 6. Percentage of profits from new categories 7. Average time to break-even 8. Change in Customer satisfaction 9. Profit growth due to new products/services 10. Percentage of profits from new products in a given period 11. Change in NPV of portfolio 12. Percentage of growth target attributable to change in product portfolio 13. Change in marketvalue or shareholder returns due to new products

mean the planning is fine but the development is ineffective and changes will have to be made in the development process.

In developing their innovations, firms need to be wary of protecting their intellectual properties so that they have the opportunity to fully capitalize on the resources invested. As part of the development process, firms need to ensure that relevant patents are filed and obtained in a timely manner (*see* INTELLECTUAL PROPERTY RIGHTS).

Many firms assess their innovation progress through the "time to market" metric. It has been argued that much of the troubles US firms have experienced with regard to innovation relates to moving projects through development and bringing them to the market in a timely manner. Being late to the market has its costs, so time to market is a critical measure of the innovation process.

Firms also need to have a balance of both big ideas and small ideas. The bolder innovations tend to be more costly to both develop and launch since there is minimal internal experience within the organization. Yet, modest changes to the existing portfolio can only grow the firm so much.

Performance outcome metrics. The vast majority of the measures used to assess innovation relates to the outcome stage. This is somewhat unfortunate because such measures can be tracked only

after the innovation is complete, not affording the opportunity to adjust the process of innovation. Yet, if the overall performance of the innovations is poor, it does imply the need to modify the process, even if it does not indicate how the process should be adjusted. The majority of measures are at this stage as these tend to be more objective than measures at earlier stages. This is also the case that the major objective is financial performance.

Perhaps, the most commonly used measure is the percentage of sales from new products over the past few years. This measure indicates how much innovations contribute to the firm's top line. Of course, an important question relating to this measure is: what constitutes a "new product?" Making slight modifications to existing products and then claiming that the innovation process is working misses the purpose of this measure and does not give a true sense of whether the innovation development process is oriented toward the future of the firm. Firms seek incremental new sales from the new products rather than merely obtain sales that replace sales from former products.

A measure which does capture whether the innovation development process is working is the success ratio of the new products. Again, this is a measure on which 100% is not the desired goal as it would indicate that the firm is not reaching far enough with the innovation efforts; yet, most companies strive to have this number be as large as is possible. The big question is what constitutes "success." This success ratio could be increased merely by setting a lower success hurdle, which would defeat the purpose of this measure.

Revenue growth, as a measure, is also frequently used by firms. Unfortunately, this measure might be achieved because of factors other than innovation, such as marketing efforts or the size of the chosen markets. Similarly, return on investments in innovations, a financial measure, is also flawed because it might reflect the impact of the level and effectiveness of the marketing efforts and not the innovation.

Perhaps, a more reasonable outcome measure is the percentage of profits from new products or profit growth from new products. This is less commonly used than percentage of sales because new products seldom make money in their early years. This measure would not reflect the true potential of new products. Hence, some companies have sought to use a measure of the percent of time to breakeven. While this measure does not reflect the innovation development process, it indicates the speed at which innovation projects get accepted in the marketplace and provide a return to the company.

Another commonly used outcome measure is overall growth in firm profits. Because the overall profit growth of the firm could be attributed to the growth decline from existing products, it would make sense for firms to focus on revenue and profit growth that are directly attributable to new products.

The most commonly used customer-related measure of innovation is customer satisfaction. This is a powerful measure for several reasons. First, it indicates whether the development process has yielded a product or service that the customer wants. It indicates the extent to which the firm has listened to the customer and responded to customer needs (*see* VOICE OF THE CUSTOMER). Second, it indicates whether further innovation development is needed. If a sufficient number of customers are not fully satisfied, there is opportunity to grow if further development is done on the product or service. More importantly, unlike other measures which indicate current market place performance, customer satisfaction is believed to be a good indicator of future performance. Customers who are satisfied are likely to be future customers as well as advocates or evangelists to others about the quality of the firm's new offering. The Net Promoter Score (NPS) is a variant of satisfaction which companies such as General Electric have adopted (Reichheld, 2006). It is based on the concept that the future performance of a product or service is driven by the number of customers who are highly satisfied and are likely to be "promoters" of the product relative to those that are less satisfied and might be "detractors." Thus, the NPS is measured as the difference between the percentage of promoters and the percentage of detractors (Reichheld, 2006).

At a firm level, an important innovation metric is firm innovativeness – the number of new offerings introduced annually by the firm. While customer satisfaction or NPS might be good

potential indicators of future success of a particular innovation, at a firm level, customer satisfaction is not significantly related to firm innovativeness (Dotzel, Shankar, and Berry, 2010). However, both innovativeness and customer satisfaction are positively related to firm value as measured by Tobin's Q (Dotzel, Shankar, and Berry, 2010). Thus, firm innovativeness may be a leading indicator of a firm's future success.

Table 2 Strategic importance of innovation[a].

	Number of metrics used
• Top strategic priority	9.5
• Second of third priority	8.0
• Not among top 3	6.8

[a]Companies use an average of 7.9 out of possible 26 metrics

Metrics Usage

All of the metrics shown in Table 1 are used, but they are not all used with the same frequency. Most firms use a stage-gate process to monitor the progress of their innovation development. Some firms use growth rate as the best proxy for assessing whether their innovation process is working. Among outcome measures, customer satisfaction is used by 44% of the firms as their primary measure of innovation success according to Boston Consulting Group (2009).

Griffin and Page (1993) conducted one of the early studies examining what measures companies were using to assess their innovation process. They found that 86% of the firms in their study use customer and financial measures and few used process measures. We reexamined this issue nearly two decades later as innovation has grown in importance. We studied the importance of innovation, the metrics used, and the perception of outcomes. With the cooperation of McKinsey & Company, we used the McKinsey Quarterly database, which consists of 18 668 executives, the majority of whom are either C-level or other senior executives. Companies represented in our sample range from quite small with revenues less than $1 million to firms with revenues in excess of $30 billion.

Not surprisingly, we found that innovation remains a top priority for the vast majority of firms. Firms use a myriad of metrics primarily to provide strategic direction for their innovation activities, guide resource allocations, and to improve innovation performance.

On average, companies use 7.9 different measures to assess their innovation process. The more strategically important they believed innovation was to the firm, the more measures they used (see Table 2).

Without question, the most commonly used measures were outcome measures rather than process or input measures (see Table 3).

The most widely used metric under input measure was number of new ideas/concepts in the new product pipeline; for process effectiveness it was the number of new products launched; and for performance outcome metric it was revenue growth.

Firms in the survey claimed that the primary use of their innovation metrics was to (in order)

Table 3 The most used measures (proportion of sample using the metric).

Inputs	*Process effectiveness*	*Performance outcomes*
• Number of new ideas/concepts in the pipeline	• Number of products launched (46%)	• Revenue growth (60%) • Customer satisfaction (48%) • Return on investment in innovation (48%) • Profit growth (45%) • Percent of sales from new product in past *N* years (44%)

Only metrics used by at least 40% of sample are shown

- assess the effectiveness of innovation spending,
- effectively allocate resources,
- improve people's accountability,
- improve overall innovation performance,
- assess progress against overall goals, and
- communicate effectively to the investment community.

Interestingly, however, we found that only 19% felt their organization's innovation metrics were effectively aligned with individual performance incentives. While the desire is to use these metrics for the above purposes, it is unclear whether the metrics used today fully accomplish any of the above objectives.

Conclusions

Firms use a large number of innovation metrics. These metrics cut across the spectrum of the innovation process but are highly skewed toward outcome measures. There clearly is no single "silver bullet" measure. While most organizations find innovation extremely important and measure it vigorously, they are still not adequately satisfied with their innovation measurement. Clearly, more work is needed in this area for us to move forward.

Bibliography

Boston Consulting Group (2009) *Innovation 2009: A BCG Senior Management Survey*, The Boston Consulting Group, Inc., Boston.

Dotzel, T., Shankar, V. and Berry, L. (2010) Service Innovativeness, Customer Satisfaction, Firm Value, and Firm Risk: Asymmetries Between Internet-enabled and Non Internet-enabled Service Innovativeness. Working Paper, Texas A&M University, TX.

Griffin, A. and Page, A.L. (1993) An interim report on measuring product development success and failure. *Journal of Product Innovation Management*, **10**, 291–308.

Philips (2009) *Philips Annual Report*, Eindoven, Netherlands.

Reichheld, F. (2006) The one number you need to grow. *Harvard Business Review*, March.

the Stage-Gate idea to launch system[1]

Robert G. Cooper

An Idea-to-Launch Roadmap

New products are critical to the survival and prosperity of the modern corporation. But product innovation is not so easy: new products fail at an alarming rate (about 1 in 10 new-product concepts succeeds); 44% of new-product projects fail to meet their profit objectives; and 49% are launched late to market (Adams and Boike, 2004; Cooper, 2001; Cooper, Edgett, and Kleinschmidt, 2003; Griffin, 1997).

A *Stage-Gate®* system is one solution that many leading companies have adopted to drive new-product projects to market quickly and effectively[2]. Almost every top-performing company has implemented a stage-and-gate system to drive their new-product projects through to commercialization, according to an APQC (American Productivity and Quality Center) benchmarking study into product innovation (Cooper, Edgett, and Kleinschmidt, 2003, 2005). PDMA (Product Development and Management Association) management best practices studies show similar adoption rates (Adams and Boike, 2004; Griffin, 1997).

A *Stage-Gate®* new-product process is simply a "playbook," "game plan," or road map to guide new-product projects from idea to launch. Here, we look at what a *Stage-Gate®* system is, and then at 10 best practices that top-performing businesses have built into their stage-and-gate methodologies. We also provide an outline of the *Stage-Gate®* system via a walk-through process.

The Structure of Stage-Gate®

The *Stage-Gate®* method breaks the product innovation process into a predetermined set of stages, each stage consisting of a set of prescribed, cross-functional, and parallel activities (see Figure 1). Gates are the entrance to each stage, and serve as quality-control and Go/Kill check points. This stage-and-gate format leads to the name "*Stage-Gate®*" process.

The *Stage-Gate®* method is based on the experiences and observations of a large number of managers and firms from research by the author and others in the field. Since *Stage-Gate®* first appeared in print, it has been implemented in hundreds of leading firms worldwide, many of which have provided an excellent setting to further improve the process (Cooper, 2008).

The stages. Stages are where the project team members execute key tasks to gather information needed to advance the project to the next gate. And there is a *fairly standard* or *prescribed list of tasks* for each stage. The specified tasks answer the following questions:

- What does the management need to know at the end of this stage to make an informed decision to move forward?
- What actions are therefore required to get this information?

For example, in Stage 2, Build the Business Case, a number of key tasks are required to deliver a business case, such as undertaking Voice-of-Customer research (*see* VOICE OF THE CUSTOMER), doing a competitive analysis (*see* COMPETITIVE ADVANTAGE), defining the product, and doing a source-of-supply assessment.

Stages are cross functional: There is no R&D or marketing stage, and handoffs from one department to another are not permitted. Rather, Stage-Gate® is an *integrated business process*, with each stage consisting of a set of parallel tasks undertaken by *people from different functional areas* within the firm, working together as a team and led by a project team leader (*see* INTEGRATED PRODUCT DEVELOPMENT).

To manage risk via a *Stage-Gate®* method, the parallel tasks in each stage are designed to gather vital information – technical, market, financial, operations – to reduce *both* the *technical* and *business risks* of the project. Each stage costs more than the preceding one, so that the process is based on incremental commitments. As uncertainties decrease, expenditures are allowed to mount, and risk is managed.

The flow of the typical *Stage-Gate®* model is shown in Figure 1:

1. *Discovery*: prework designed to discover opportunities and to generate new-product ideas (*see also* IDEA MANAGEMENT).

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

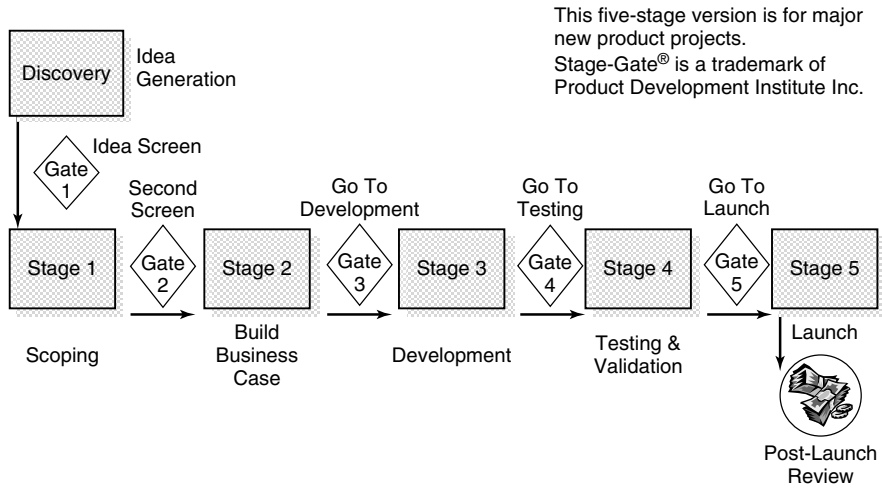


Figure 1 The five-stage Stage-Gate® process is used by many businesses to drive new products from idea to launch.

2. *Scoping*: a quick, preliminary investigation, and scoping of the project. This stage provides inexpensive information – based largely on desk research – to enable the list of projects to be narrowed before Stage 2.
3. *Build the business case*: a much more detailed investigation involving primary research – both market and technical – leading to a *business case*. This is where the bulk of the vital homework is done, which results in a business case: the product definition, the project justification, and a project plan (*see* CONCEPT TESTING).
4. *Development*: the detailed design and development of the new product, along with some product testing work (*see* PRODUCT TESTING). The deliverable at the end of Stage 3 is an alpha-tested or lab-tested product. Full production and market launch plans are also developed in this potentially lengthy stage.
5. *Testing and validation*: tests or trials in the marketplace, lab, and plant to validate the proposed new product, and its marketing and production or operations. Tasks include field trails or beta tests; test market or trial sell; and operations trials (*see* PRETEST MARKET MODELS).
6. *Launch*: commercialization – beginning of full operations or production, marketing, and selling. Here the market launch, operations, distribution, and postlaunch plans are executed (*see* LAUNCH STRATEGIES).

The gates. Preceding each stage is an entry gate or a Go/Kill decision point. Effective gates are central to the success of a fast-paced, new-product process:

- Gates serve as quality-control checkpoints, ensuring that the project is executed properly.
- Gates also serve as Go/Kill and prioritization decisions points: mediocre projects are culled out at each successive gate.
- Finally, gates are where the action plan for the next stage is agreed, and the resources needed to execute the plan are committed.

Gate meetings are usually staffed by senior managers from different functions – the gatekeepers – who own the resources required by the project team for the next stage.

Gates have a common format:

1. *A set of required deliverables*: what the project team must bring to the gate decision point. These deliverables are visible, are based on a standard menu for each gate,

and are decided at the output of the previous gate. Management's expectations for project teams are thus made very clear.
2. *Criteria*: the project is judged against these to make the Go/Kill and prioritization decisions.
3. *Defined outputs*: for example, a decision (*Go, Kill, Hold or Recycle*), an approved action plan for the next stage including resource requirements, and a list of deliverables for the next gate.

Types of gate criteria. Each gate has it own set of criteria for use by the gatekeepers:

- *Readiness check*: These are Yes/No questions that check whether the key tasks have been completed, and that all deliverables are in place for that gate. A "No" signals a recycle to the previous stage because the project is *not ready* to move on.
- *Must meet*: These are Yes/No or "knock-out" questions that include the minimum business criteria that a project must meet to move forward. A single consensus "No" signals a Kill decision.
- *Should meet*: These are highly desirable project characteristics that are used to distinguish between superb projects and the minimally acceptable ones. These are typically in a scorecard format (*see* PORTFOLIO MANAGEMENT) (Cooper, Edgett, and Kleinschmidt, 2002).

Building in best practices. A number of best practices are built into the *Stage-Gate®* idea-to-launch system to yield superlative results (*see* SUCCESS FACTORS FOR NEW-PRODUCT DEVELOPMENT):

Tough gates with teeth for sharper focus, better project prioritization. Most businesses' new-product efforts suffer from a lack of focus: too many projects, and not enough resources to execute them well (Cooper, 2005; Cooper, Edgett, and Kleinschmidt, 2002). Adequate resources is a principal driver of businesses' new-product performance; but a lack of resources plagues too many development efforts (Cooper, Edgett, and Kleinschmidt, 2003).

The need is for a *new-product funnel* (rather than *tunnel*) that builds in tough Go/Kill decision-points in the form of gates; poor projects are weeded out and more focus is the result. The expectation is that a significant percentage of projects will be killed at each gate, especially at the earlier gates (Gates 1–3 in Figure 1). The gates thus become the *quality-control check points* in the new-product process, checking the quality, merit, and progress of the project.

Products with competitive advantage – differentiated products, unique benefits, superior value for the customer. Product superiority is one key to new-product success, yet all too often, when applying a new-product process, there is no attempt to seek truly superior products (*see also* COMPETITIVE ADVANTAGE; VALUE PROPOSITION). An effective Stage-Gate® system builds in the quest for product advantage:

- Some of the criteria at every gate focus on product superiority. Questions such as "Does the product have at least one element of competitive advantage?" become vital questions to rate projects.
- Customer-related tasks designed to deliver product superiority are included in each stage of the process. Examples are given in the next section.
- Project teams are required to deliver evidence of product superiority to gate reviews.

A strong market orientation with Voice-of-Customer inputs throughout. A strong market orientation – executing the key marketing activities in a quality fashion – must be built into the new-product process. Six best-practice marketing actions that are incorporated into the stages of a robust Stage-Gate® system include the following:

- customer-focused ideation to gain insights into customer problems;
- preliminary market assessment in the very early phases of the new product project;
- Voice-of-Customer research to identify unmet or unarticulated needs;

- competitive product analysis to determine their strategy and how to surpass them;
- value-in-use analysis to gauge the economic value of the product to the customer;
- concept tests, preference tests, and trial sells to validate the product and the project.

Front-end loading projects leading to sharp, early, and stable product definition. New-product success or failure is largely decided in the first third of the project – in those crucial steps and tasks that precede the actual development of the product. The up-front homework defines the product and builds the business case for development. The ideal new-product process ensures that these early stages are carried out and that the product is fully defined before the project is allowed to become a full development project.

The need for solid up-front homework parallels the case for a strong market orientation. Top performer businesses ensure that the new-product process includes solid homework (Stages 1 and 2 in Figure 1) and stable, fact-based product definition. For example, they build in a product-definition check-point at Gate 3; and they halt projects if the homework and product definition are not in place.

A fast-paced game plan via parallel processing and simultaneous execution. New-product teams face a dilemma. They are urged by senior management to compress the cycle time but also to cut down the failure rate – do it right! (*see also* ACCELERATED PRODUCT DEVELOPMENT; INTEGRATED PRODUCT DEVELOPMENT) *Simultaneous execution* is one solution to the need for a complete and quality process, yet a process that meets time pressures (Cooper and Edgett, 2005; Morgan, 2005). Traditionally, new-product projects have been managed via a *series approach*: one task strung out after another, in sequence. Phrases such as "hand off" or even "throwing it over the wall" are common in this *relay-race approach*.

In contrast, with simultaneous execution, many activities are undertaken *in parallel* and *concurrently* rather than in series. The process is far more intense than a relay race, with more work getting done in an elapsed time period. Moreover, there is less chance of a task being skipped because of lack of time, as each task is done in parallel, not in series, and hence does not extend the critical path.

A true cross-functional team approach. The multifunctional nature of innovation coupled with the desire for simultaneous execution means that a *cross-functional team approach* is mandatory (*see also* CROSS-FUNCTIONAL TEAM). The essential ingredients of this approach are

- a cross-functional project with committed team players from the different functional areas;
- a defined team leader, driving the entire project beginning to end, and with formal authority (coopting authority from the functional heads);
- a fluid team structure, with new members being brought in or dropped as work requirements demand;
- only a small core group of responsible and committed team players from beginning to end;
- most important, a team that is accountable for the entire project's end results (not just responsible for one part of the project).

Adaptable and agile via spiral development. Spiral development is one way that fast-paced companies cope with changing or fluid information as the project proceeds (Cooper and Edgett, 2005). Often, a project team charges into development using a product definition based on information that was right at the time. But it was not right, or customer requirements changed, and when the product is ready for testing or market launch (Stage 4 or 5 in Figure 1), it is discovered that the product is not quite right. This *traditional linear approach* means that the team must then recycle back to the development stage, and make the necessary changes to the product and its design.

By contrast, smart project teams practice *spiral development* and adapt and adjust their project over time, as in Figure 2. Like the "linear team" above, they do their front-end homework (Voice-of-Customer work, competitive analysis,

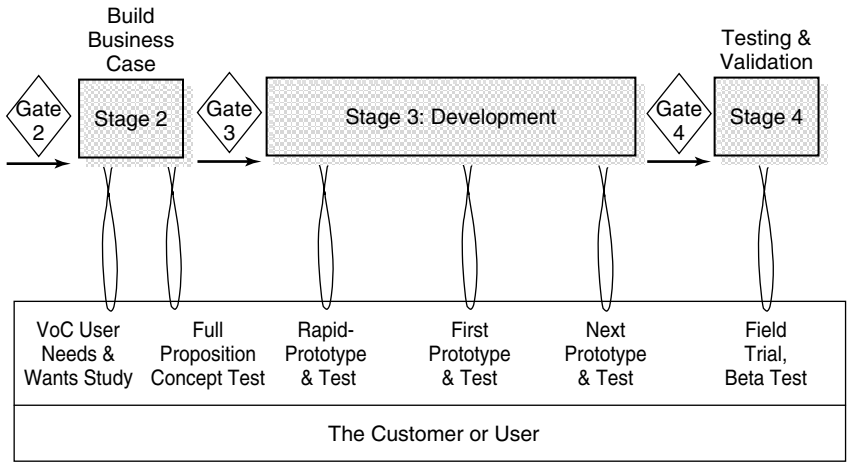


Figure 2 Stage-Gate® builds in adaptability and agility via a series of "build–test–feedback–revise" iterations or loops.

and concept testing) and define the product based on best-available information. But quickly, the "adaptive-spiral team" creates the first version of the product (often a virtual one) and tests it with the customer, seeking immediate feedback. The team uses this feedback to produce the next, more complete version of the product – a working model or protocept. In this way, these fast-paced teams remove unnecessary rework and quickly move to the finalized product by undertaking a series of these iterative steps or loops: *build, test, obtain feedback, and revise*.

Risk-adjusted and scalable. When originally conceived, companies typically designed a *single* stage-and-gate process – "one size fits all projects," usually designed for *the most complex development projects* (Cooper, 2008). Most projects, however, were much simpler – modifications and improvements – and so teams simply circumvented the "large process." The dilemma was that these small projects consumed the bulk of development resources, and so the firm ended up having most of the funds spent on projects that were "outside the system."

Now there are streamlined versions of Stage-Gate®, including Lite and XPress, designed to handle lower risk, simpler projects. For example, in the model in Figure 3, all proposed development projects enter Gate 1 on the left for an initial screen and routing decision, and depending on the risk level, follow different versions of Stage-Gate®. Additionally, some companies engaged in very innovative developments – for example, advanced technology projects – have adopted quite different development processes to handle these high-risk, longer-term, and highly uncertain projects (Cooper, 2006).

A lean, efficient process with time wasters removed. The idea-to-launch system must be built for speed. This means eliminating all the time wasters and work that add no value in the current new-product process. Using principles borrowed from lean manufacturing, management undertakes "value stream analysis," reviewing the development process from end to end, scrutinizing every required procedure, meeting, or paperwork that must completed (Kennedy, 2003), and removing any work that does not add value. The end result is that current versions of Stage-Gate® are much leaner and more efficient than previous systems.

Performance metrics in place. The idea-to-launch system should feature solid performance metrics, so that senior management can assess how well the development process is working and hold project teams accountable for results. For individual new-product projects,

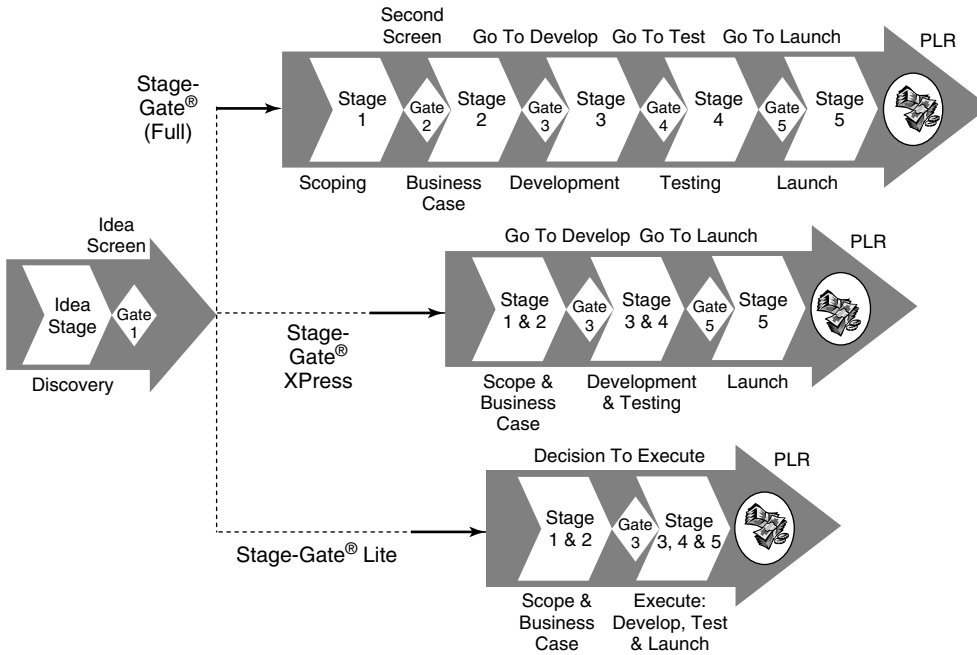


Figure 3 Stage-Gate® XPress and Lite are used for lower-risk projects. Major new-product projects go through the full five-stage process (top). Moderate-risk projects, including extensions, modification, and improvements, use the XPress version (middle). Sales-force requests (very minor changes) use the Lite process (bottom).

success metrics often include (*see also* INNOVATION METRICS) the following:

- first-year sales (versus the sales forecast in the business case at Gate 3);
- product profitability versus that forecast in the business case;
- on-time performance (actual versus promised launch date) (Cooper, Edgett, and Kleinschmidt, (2003, 2005)).

Top-performing companies also build in a *postlaunch review* 12–18 months after launch, where these metrics are used to gauge the *ultimate success* of the project. Here the project's actual results are assessed versus those results promised when the project was approved at Gate 3. Additionally, sales, profits, and on-time performance results for individual projects can be aggregated or averaged to yield performance metrics *for the business's entire new product effort*.

A Walk-through of the Stage-Gate® System

Here is a more detailed look at the *Stage-Gate®* system – of what is involved at each stage and gate in the model outlined in Figure 1 (Cooper, 2001, 2005, 2008).

Discovery stage. Ideas are the vital feedstock to the process. Many companies consider ideation so important that they handle this as a formal stage in the process, called *Discovery*, with many activities built into this stage, such as (Cooper and Edgett, 2008)

- undertaking fundamental technical research, seeking new technological possibilities (Cooper, 2006);
- doing Voice-of-Customer research to uncover unmet, unspoken needs (includes ethnography, working with lead users, and depth interview with customers);

- competitive analysis and reverse brainstorming competitive products;
- installing an idea suggestion scheme to solicit ideas from the company's own employees;
- using strategic planning to uncover disruptions, gaps, and opportunities;
- open innovation – welcoming ideas and solutions from outside the company (*see* OPEN INNOVATION).

Gate 1: Idea Screen. Idea screening is the first decision to commit resources to the project: If the decision is Go, the project moves into the scoping or preliminary investigation stage. Gate 1 signals a preliminary but tentative commitment to the project, and is a gentle screen that amounts to subjecting the project to a handful of key must-meet and should-meet criteria. Financial criteria are typically not part of this first screen, since relatively little reliable financial data are available here.

Stage 1: Scoping. This first and inexpensive homework stage determines the project's technical and marketplace merits: a quick scoping of the project, involving only desk research (no primary research). Stage 1 is often done in less than one calendar month's elapsed time, and 5-10 person-days' work effort. The key tasks include

- *a preliminary market assessment* whose purpose is to determine market size, market potential, and likely market acceptance, and to begin to shape the product concept;
- *a preliminary technical assessment* whose purpose is to assess development and manufacturing routes (or source of supply), technical and operations feasibility, time and cost to execute, and technical, legal, and regulatory risks.

Stage 1 thus provides for the gathering of both market and technical information – inexpensively and quickly – to enable a cursory financial and business analysis as input to Gate 2.

Gate 2: Second Screen. The project next proceeds to more rigorous screen at Gate 2, where the project is reevaluated with new information obtained in Stage 1. If the decision is Go, the project moves into a heavier spending stage. A checklist and scoring model facilitate this gate decision.

Stage 2: Build the Business Case. Stage 2 is a detailed investigation stage, which defines the product and verifies the attractiveness of the project prior to heavy spending. It is also the *critical homework* stage, the one found to be so often weakly handled.

Stage 2 sees Voice-of-Customer research undertaken to exactly determine the customer's needs, wants, and preferences. Competitive analysis is also a part of this stage. Another market activity is concept testing where potential customers' reactions to the concept are gauged.

The detailed *technical appraisal* at Stage 2 focuses on the technical feasibility of the project. Here, customer needs are translated into a technically and economically feasible solution on paper – preliminary design or laboratory work, but not a full development effort. An operations appraisal is undertaken, where issues of manufacturability, source of supply, costs, and investment required are investigated. And detailed legal, patent, and regulatory assessment work is done to remove risks and map out the required actions (*see also* INTELLECTUAL PROPERTY RIGHTS).

Finally, a detailed *business and financial analysis is* conducted, typically a net-present-value calculation, complete with sensitivity analysis to consider possible risks. The result of Stage 2 is a *business case* for the project: the *product definition* is agreed upon; and a thorough *project justification* and *detailed project plan* are developed.

Gate 3: Go to Development. This is the final gate prior to the development stage, the last point at which the project can be killed before entering heavy spending. Once past Gate 3, financial commitments are substantial, thus Gate 3 is usually staffed by senior management.

Gate 3 subjects the project to a rigorous set of readiness-check, must-meet, and should-meet criteria. And because a heavy spending commitment is the outcome of a Go decision, the results of the financial analysis are an important part of this screen.

Stage 3: Development. Stage 3 sees the implementation of the development plan and the

physical development of the product. Lab tests, in-house tests, or alpha tests ensure that the product meets requirements under controlled conditions. Also, the operations or source-of-supply process is mapped out. Extensive in-house tests, alpha tests, or lab tests usually take place in this stage as well. The deliverable at the end of Stage 3 is a lab-tested or alpha prototype of the product.

The emphasis in Stage 3 is on technical work. But marketing and operations activities also proceed in parallel. For example, market-analysis and customer-feedback work continue concurrently with the technical development, with constant customer feedback sought on the product as it takes shape via *spiral development*. Meanwhile, detailed test plans, market launch plans, and production or operations plans, including operations facilities requirements, are developed. An updated financial analysis is prepared, while regulatory, legal, and patent issues are resolved.

Gate 4: Go to Testing. This postdevelopment gate is a check on the progress and the continued attractiveness of the project. Development work is reviewed, ensuring that the developed product is indeed consistent with the original definition specified at Gate 3. This gate also revisits the project's economics via a revised financial analysis. The validation plans for the next stage are approved, and the detailed market launch and operations plans are reviewed for future execution.

Stage 4: Testing and Validation. This stage validates the viability of the project: the product, its production process, customer acceptance, and its economics. The key tasks are

- *In-house product tests*: extended lab tests or alpha tests to check on product quality and product performance under controlled or lab conditions.
- *User, preference, or field trials of the product*: to verify that the product functions under actual use conditions, and also to measure purchase intent.
- *Trial, limited, or pilot operations*: to test, debug, and prove the operations process, and to determine more precise costs and throughputs.
- *Pretest market, test market, or trial sell*: to gauge customer reaction, measure the effectiveness of the launch plan, and determine expected market share and revenues.

Gate 5: Go to Launch. This final gate opens the door to full commercialization: market launch, and operations start-up. Criteria for passing the gate focus largely on the expected financial return, the project's readiness for launch, and the appropriateness of the launch and operations start-up plans.

Stage 5: Launch. This final stage involves implementation of both the market launch plan and the operations plan. Equipment is acquired, installed, and commissioned (although sometimes this is done earlier in Stage 4, as part of the Stage 4 operations trials); the logistics pipeline is filled; and selling begins.

Post-Launch Review. Two Post-Launch Reviews are typical. The first, an interim review, occurs about two to four months after launch when initial commercial results are available. Here a *retrospective analysis* is done, which assesses the project's strengths and weaknesses, and identifies what can be learned from the project. This task builds in *continuous improvement*.

The final review is held once the project's commercial results are known, typically 12–18 months after launch. Here, the project and product's performance is reviewed: The latest data on the revenues, costs, expenditures, profits, and timing are compared to projections made at Gates 3 and 5 to gauge performance. This final review marks the end of the project.

Conclusion

Product innovation is perhaps the most important endeavor of the modern corporation. Without a systematic new-product process, however, often product development suffers. The *Stage-Gate*® system is an enabler or guide, building in best practices and ensuring that key tasks and decisions are undertaken better and faster. But *Stage-Gate*® is considerably more complex than the simple diagram in Figure 1

suggests; there are many intricacies in the details – both the "what's" and the "how to's." And implementing the process is also a major challenge. Many leading companies, however, have adopted a world-class idea-to-launch *Stage-Gate*® system, and the results have been positive: better, faster, and more profitable new-product developments.

Endnotes

[1] This material is based on several previous publications by the author. See Adams and Boike (2004) and Cooper (2005).

[2] The term *Stage-Gate*® was coined by the author, and is a trademark of the Product Development Institute Inc.: www.prod-dev.com.

Bibliography

Adams, M. and Boike, D. (2004) PDMA foundation CPAS study reveals new trends. *Visions*, **XXVIII** (3), 26–29.

Cooper, R.G. (2001) *Winning at New Products: Accelerating the Process from Idea to Launch*, 3rd edn, Perseus Books, Reading.

Cooper, R.G. (2005) *Product Leadership: Pathways to Profitable Innovation*, 2nd edn, Perseus Books, Reading.

Cooper, R.G. (2006) Managing technology development projects – different than traditional development projects. *Research-Technology Management*, **49** (6), 23–31.

Cooper, R.G. (2008) The Stage-Gate idea-to-launch process – update, what's new and NexGen systems. *Journal of the Product Innovation Management*, **25** (3), 213–232.

Cooper, R.G. and Edgett, S.J. (2005) *Lean, Rapid and Profitable New Product Development*, Product Development Institute, Ancaster, www.stage-gate.com.

Cooper, R.G. and Edgett, S.J. (2008) Ideation for product innovation: what are the best sources? *Visions*, **XXXII** (1), 12–17.

Cooper, R.G., Edgett, S.J., and Kleinschmidt, E.J. (2002) *Portfolio Management for New Products*, 2nd edn, Perseus Publishing, New York.

Cooper, R.G., Edgett, S.J., and Kleinschmidt, E.J. (2003) *Best Practices in Product Innovation: What Distinguishes Top Performers*, Product Development Institute, Ancaster, www.prod-dev.com.

Cooper, R.G., Edgett, S.J., and Kleinschmidt, E.J. (2005) Benchmarking best NPD practices – part III: the NPD process and decisive idea-to-launch practices. *Research-Technology Management*, **47** (6), 43–55.

Griffin, A. (1997) Drivers of NPD Success: The 1997 PDMA Report, Product Development & Management Association, Chicago.

Kennedy, M.N. (2003) *Product Development for the Lean Enterprise*, The Oakley Press, Richard.

Morgan, J. (2005) Applying Lean Principles to Product Development, SAE International Society of Mechanical Engineers, www.shop.sae.org.

diffusion of innovation

Deepa Chandrasekaran and Gerard J. Tellis

Definitions

An innovation is an idea, practice, or object that is perceived as being new by an individual or other unit of adoption (Rogers, 1995). There are several definitions of "diffusion." Some scholars define *diffusion* as a special type of communication where the message refers to an innovation (Rogers, 1995; Mahajan, Muller, and Wind, 2000). In this sense, diffusion refers to the process by which an innovation is communicated through certain channels over time among the members of a social system. Other scholars differentiate the *phenomenon*, diffusion, from its *drivers*, such as product's price, advertising, or word-of-mouth communication. In this sense, diffusion refers to the spread of an innovation across social groups or markets over time (Stoneman, 2002; Chandrasekaran and Tellis, 2007).

Shape of the Diffusion Curve

Numerous studies in a variety of disciplines suggest that the typical plot of cumulative adoptions of new products or technologies over time is an S-shaped curve (e.g., Mahajan, Muller, and Wind, 2000, *see also* TECHNOLOGY S-CURVE). The Bass diffusion model (Bass, 1969) was used to depict the first purchase growth of a new durable product (*see* BASS MODEL). The Bass model suggests that mass media and interpersonal communications are important in determining the speed and shape of the diffusion processes (Mahajan, Muller, and Wind, 2000). The Bass model has three key parameters: the coefficient of innovation or external influence (p), the coefficient of imitation or internal influence (q), and the market potential (m). The diffusion curve is S-shaped when $q > p$, and is more pronouncedly so as the q/p ratio increases (Van den Bulte and Stremersch, 2004). Recent interpretations (Mahajan, Muller, and Peres, 2008) regard the internal coefficient q to represent all consumer interdependencies, including signals, externalities, and interpersonal communications. Interpersonal communications take place when consumers actively engage in collecting and processing information from previous adopters. Consumers may also be influenced by nonpersonal information, or signals from the market such as information on the number of prior adopters. The utility of the innovation may also increase as more consumers adopt a product, and hence network effects may influence adoption.

The diffusion curve is related to the concept of the PRODUCT LIFE CYCLE (PLC) (Golder and Tellis, 2004). New products often see an initial phase of low sales (introduction), followed by a period of rapid increase in sales (growth), followed by a period of slowdown in sales (maturity), and a period of steady decline in sales (decline). Recent research has documented two distinct transition points – takeoff, marking the transition from the introduction to the growth stage of the PLC (*see* TAKEOFF), and slowdown, marking the transition from the growth to the maturity stage of the PLC (Agarwal and Bayus, 2002; Goldenberg, Libai, and Muller, 2002; Golder and Tellis, 2004).

Drivers of Diffusion

Apart from consumer interdependencies, the diffusion of innovations is broadly thought to be influenced by three factors: the characteristics of the innovation, the characteristics of the adopters, and firm strategy.

Characteristics of the innovation. The specific characteristics of innovations may influence their diffusion. Rogers (1995) outlines five characteristics of innovations that may influence the rate of adoption – *relative advantage* or the degree to which an innovation is better than the idea it supersedes; *compatibility* or the extent to which an innovation is perceived to be consistent with a consumer's past experiences, values, and lifestyle; *complexity* or how difficult it is for a consumer to learn to use the new innovation; *trialability* or the extent to which the product can be tried on a limited basis; and *observability* or whether the benefits of this new product can be easily explained and demonstrated.

Researchers on diffusion of durable goods have observed that information or entertainment products (referred to as *brown goods*, *consumer*

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

electronics items) diffuse faster than work products (referred to as *white goods*, *home appliances*), possibly because they are more visible, used by many more members of a household, and are often status symbols (Chandrasekaran and Tellis, 2007).

Characteristics of adopters. Diffusion scholars have identified certain adopter categories on the basis of innovativeness, or how early individuals in a social system adopted an innovation. On the basis of the noncumulative distribution of adopters, which takes the form of a bell-shaped curve, Rogers (1995) distinguished five categories of adopters, which include the innovators (2.5% of adopters), early adopters (13.5%), early majority (34%), late majority (34%), and laggards (16%). Mahajan, Muller, and Srivastava (1990) derive this classification from the Bass diffusion model with innovators (0.2–2.8% of adopters), early adopters (9.5–20%), early majority (29.1–32.1%), late majority (29.1–32.1%), and laggards (21.4–23.5%).

The innovators and early adopters (the early market) are often the target of a firm's marketing efforts, since it is believed that they are more innovative, are often the opinion leaders, and hence influence the actions of later adopters. Recent research, however, questions this approach arguing instead that the social contagion process is broken at the point of transition from the early market to the late market because of a lack of similarity between, and communication across, these adopter segments (Goldenberg, Libai, and Muller, 2002). There is a need to research when such discontinuities appear, what marketing strategies (such as price and promotion) can enable a smooth transition across segments, and whether companies should directly target the main market (Mahajan, Muller, and Wind, 2000).

Firm Strategy. Firm technological choices, price declines, advertising, distribution, and firm entry choices may also have a big impact on the takeoff and diffusion of a new product (Agarwal and Bayus, 2002; Golder and Tellis, 1997; Kuester, Gatignon, and Robertson, 2000). There is, however, little consensus on what strategic factor dominates in driving diffusion, for instance, on the role of price declines versus firm entry in driving takeoff.

Models of Diffusion

Over the past 35 years, a vast body of literature has sought to enrich the basic Bass model by including marketing variables, supply restrictions, multiproduct interactions, incorporating time-varying parameters, replacement purchases, multiple purchases, and trial and repeat purchases, and improved estimation techniques (see, for instance, Chandrasekaran and Tellis (2007) and Mahajan, Muller, and Peres (2008) for summaries and references). This body of research indicates that more realistic models combined with improved estimation techniques can lead to the development of better forecasts of peak sales and the sales evolution during the growth stage.

Global Diffusion

Much marketing research relates to the spread of innovations across nations (*see also* GLOBAL PRODUCT DEVELOPMENT). Marketing research has focused on the nature of global diffusion processes, differences across countries in terms of innovativeness in adoption of new products, drivers of global diffusion, and market-entry strategies. There is an emerging consensus (Chandrasekaran and Tellis, 2007; Dekimpe, Parker, and Sarvary, 2000) that product diffusion is influenced by both *within-country* factors such as the wealth of a country, and cultural dimensions (such as social system heterogeneity, cosmopolitanism, individualism/collectivism) and *cross-country* factors (cross-national learning or lead–lag effects).

There is a great need to reach a consensus on operationalizations of the covariates used in international studies, and to expand the product context to beyond the study of durable goods. A key research question in diffusion is whether there has been acceleration in the *speed* of diffusion, and what are the drivers that lead to such an increase in the speed of diffusion. Recent studies find some evidence for a systematic increase in diffusion speed over time (Van den Bulte, 2000). In the context of rapid globalization, the research

challenge remains to document whether diffusion speed is increasing across contexts, what the drivers of the increase in diffusion speed might be, and whether differences in consumer adoption of new products across countries are also declining with time.

Bibliography

Agarwal, R. and Bayus, B.L. (2002) Market evolution and sales takeoff of product innovations. *Management Science*, **48** (8), 1024–1041.

Bass, F.M. (1969) A new product growth model for consumer durables. *Management Science*, **15** (5), 215–227.

Chandrasekaran, D. and Tellis, G.J. (2007) A critical review of marketing research on diffusion of new products, in *Review of Marketing Research*, vol. 3 (ed. N.K., Malhotra), M.E. Sharpe, Armonk, pp. 39–80.

Dekimpe, M.G., Parker, P.M., and Sarvary, M. (2000) Multimarket and global diffusion, in *New Product Diffusion Models* (eds V. Mahajan E. Muller, and Y. Wind), Kluwer Academic Publishers, Boston.

Goldenberg, J., Libai, B., and Muller, E. (2002) Riding the saddle: how cross-market communications can create a major slump in sales. *Journal of Marketing*, **66** (2), 1–16.

Golder, P.N. and Tellis, G.J. (1997) Will it ever fly? Modeling the takeoff of really new consumer durables. *Marketing Science*, **16** (3), 256–270.

Golder, P.N. and Tellis, G.J. (2004) Going, going, gone: cascades, diffusion, and turning points of the product life cycle. *Marketing Science*, **23** (2), 207–218.

Keuster, S., Gatignon, H., and Robertson, T.S. (2000) Firm strategy and speed of diffusion, in *New Product Diffusion Models* (eds V. Mahajan, E. Muller, and Y. Wind), Kluwer Academic Publishers, Boston.

Mahajan, V., Muller, E., and Peres, R. (2008) Innovation Diffusion and New Product Growth: Beyond a Theory of Communications, working paper.

Mahajan, V., Muller, E., and Srivastava, R.K. (1990) *Journal of Marketing Research*, Determination of adopter categories using innovation diffusion models **27** (1), 37–50.

Mahajan, V., Muller, E., and Wind, Y. (2000) New product diffusion models: from theory to practice, in *New Product Diffusion Models* (eds V. Mahajan, E. Muller, and Y. Wind), Kluwer Academic Publishers, Boston.

Rogers, E. (1995) *Diffusion of Innovations*, The Free Press, New York.

Stoneman, P. (2002) *The Economics of Technological Diffusion*, Blackwell, Cambridge.

Van den Bulte, C. (2000) New product diffusion acceleration: measurement and analysis. *Marketing Science*, **19** (4), 366–380.

Van den Bulte, C. and Stremersch, S. (2004) Social contagion and income heterogeneity in new product diffusion: a meta-analytic test. *Marketing Science*, **23** (4), 530–544.

service innovation management

Helen M. Perks

Definitions

Service innovation management has emerged as a distinct area of practice and academic study since the 1980s. However, compared to the management of "product" innovation, research on service innovation management remains limited and fragmented. A comprehensive review of innovation research published between 1989 and 2004 found that only 6% of the articles in the dataset included services and even fewer papers focused on service innovation explicitly. Compared to the two previous decades, service innovation research has grown rapidly since 2000. However, it is still underrepresented, accounting for only 5% of product development research publications between 2000 and 2004 (Page and Schirr, 2008).

There is a high degree of ambiguity regarding the term *service innovation*. It is often used interchangeably with the construct *new-service development*. Service development has its roots in marketing and service quality, whereas service innovation emanates from economics and business strategy research. The latter body of research is traditionally concerned with theories of entrepreneurship and technological development, whereas the former is concerned with exploring service-development practices and their management. The management of service innovation incorporates the strategic nature of developing and offering new services along with the management of the innovation process itself.

Definitional debates, within both the service innovation and development areas, center on the concept of newness. Such debates are conceptual in nature and are largely based on anecdotal evidence. A classic definition of service innovation (Johnson *et al.*, 2000) concentrates on the nature of the service outcomes. From this perspective, innovation in new services can range from simple incremental-style changes to major radical innovations (*see* RADICAL INNOVATION). More comprehensively, service innovation definitions have incorporated procedural or operational aspects of services. Synthesizing such debates, service innovation can be categorized in two ways:

- *Service offering innovation*: This is an overarching term incorporating incremental improvements and radical changes to service-performance attributes. Part of the infamous Booz *et al.* (1984) marketing-oriented categorization is frequently applied to services, indicating varying types of newness within the service offering, from both the supplier and customer perspective. Such notions of highly new (breakthrough or radical) or not-so-new (derivative or incremental) focus on the new service offering. The concept of development in the service attributes can also be extended to include the augmented service offer. Here innovation centers on expanding the basic service offering. This is usually conceived as innovation through forms of support (such as delivery, distribution, promotion, expertise, and customization). These definitions take an external final offering perspective. There is no categorical description or indication of the internal development resources or competencies linked to the service innovation (Menor, Tatikonda, and Sampson, 2002).
- *Service process innovation*: The cost and repositioning definitions suggested by Booz *et al.* (1984) encompass innovation in the process, rather than the offering itself. Process innovation within services concerns alterations to the business processes within the organization, such as the introduction of new technology or changes in working practices. Such a definition incorporates newness in the service concept, namely, how the service is offered. Hence this definition provides an indication of the degree of change from existing systems and routines, suggesting levels of internal newness.

The lack of definitional consensus creates problems in empirically researching service innovation. New services are often treated in aggregate, and clarity in definition, prior to investigations, is problematic. The scope of service offerings is also broad and diverse. Service innovation primarily occurs in

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

sectors such as banking and finance, leisure and hospitality, travel services, telecommunications services, information services, facilities management, education, professional /consulting services, and health care. Most academic research has been conducted in the financial-services sector, although high-technology services are receiving recent attention. Services may be offered to end consumers, businesses, or both.

Differences between Product and Service Innovation

The management of service innovation is frequently compared and contrasted with that of product innovation. Services differ from products in four fundamental ways and this has major consequences for service innovation management (Fitzsimmons and Fitzsimmons, 2001). First, the output of the service innovation process, namely, the service offering, is inherently a process rather than a tangible entity (although efforts are commonly made to make services less abstract and incorporate physical productlike attributes). As a result, traditional activities of the product innovation process, such as concept development and testing, are difficult to carry out. Further, absence of a tangible goal or outcome can impede a common vision for participants in the service innovation process. This can lead to inconsistency and misinterpretations in the development effort. As services lack the physical attributes of products, ad hoc modifications can be easily made by front-line staff in the delivery of the service. Research has shown that this can result in variable customer-service quality and poor company reputation Hence, mechanisms to structure or model the service delivery operations are advocated. Further, such intangibility means service innovations are difficult to patent and susceptible to copying by competitors.

Secondly, the heterogeneity of services creates difficulties in standardizing the service experience. Variation at the delivery/consumption point limits the extent to which service innovation can be controlled or managed. It also helps explain why aspects of the service innovation management (such as emphasis on employee training and customer-management skills) differ from product innovation. As with intangibility, the heterogeneous nature of services complicates implementation of traditional innovation activities, such as prototype or final-product testing with customers (*see* PROTOTYPE). Thirdly, the simultaneous act of production and consumption of services (simultaneity), and fourthly, the perishable nature of services, exacerbate problems of matching of supply, prevent the ability to hold stock, and put much emphasis on capacity planning for service providers. However, the close interaction of the customer with service providers in producing the service has also led to increased opportunities for customer involvement in innovation and service design.

The Service Innovation Process

In essence, service innovation is a difficult and messy activity to manage. Substantive debate, in service innovation management research, centers on the nature of the innovation process in services. Innovation investigators are concerned with whether service organizations employ a formal or informal process in service innovation. Many academics suggest that few service providers have a formalized process (with specified stages or activities) in place. Some take the view that service innovation occurs because of intuition, flair, or luck.

Despite this, prominent scholars advocate the adoption of a systematic process and suggest it contributes to new-service success (de Brentani, 1995; Alam and Perry, 2002). Empirical work researching such structured service innovation processes has predominantly focused on identifying activities that concur with models offered for new-product innovation. These often take a sequenced and staged approach, although it is debatable whether activities occur concurrently or in sequential order. Indeed, a body of research emphasizes the iterative, cyclic, and nonsequential nature of new-service innovation (Johnson *et al.*, 2000; Menor, Tatikonda, and Sampson, 2002). Booz *et al.* 1984 suggested a six-stage model of strategy, idea generation, screening and evaluation, business analysis, development, testing, and commercialization that is frequently adopted in service innovation process research. Such formalized systems can provide rules and

procedures to help structure and guide the development process. Other research offers augmentations pertinent to services. Such new extensions enrich our understanding of the conditions and the environment that support or facilitate the service innovation process. The widely used model suggested by Johnson *et al.* (2000), for example, depicts a four-stage service innovation process, with 13 specific tasks, and identifies the components of the organization that are involved in the process.

As in product innovation, service researchers frequently adopt a staged or time-based approach to understanding or exploring the service innovation process. Such work investigates the core activities carried out at each stage. These are often referred to as *early*, *mid-*, and *late-stage activities*. Early stage activities in service innovation are equivalent to the fuzzy front end in product innovation (*see* FRONT END OF INNOVATION). They concern the creation and analysis of ideas and opportunities (*see* OPPORTUNITY IDENTIFICATION). Idea generation in service organizations is perceived as relatively easy. This perception is linked to the dominant nature of innovation in services, namely, incremental improvements and minor modifications to the service offering. Empirical research has shown that few ideas, in service organizations, are radical and most are copied from competitors, who are identified as the most important source (Johne and Storey, 1998). Not surprisingly, idea generation is found to be undertaken on an ad hoc basis (Easingwood, 1986). The high degree of service copying can also explain a low level of market-research use in service innovation. Further, the lack of prototyping and difficulties in reproducing the in-use service experience make market research problematic. However, academic studies suggest that service managers' perceptions of customer needs are frequently incorrect. Similarly, concept testing is difficult to implement. The challenges of precisely describing or demonstrating service benefits, at the concept stage, are acute, although recent advances in simulation techniques and technology go some way to meeting these challenges. Studies have also indicated that evaluation or screening of service ideas and concepts is usually informal. Beyond traditional financial measures, little is known about the specific screening criteria employed. However, some research has drawn attention to service organizations' assessment of the potential impact of the new service offering on the image of the service organization. Brand image can act as a protective mechanism against competitive emulation. Hence organisations should ensure that service innovation aligns with a strong reputation.

In the development stage or activity set of service innovation, research tends to adopt two perspectives: the customer-facing service concept and the operations or delivery concept. The first domain is classically conceptualized as *front office* where service staff interact with customers. The latter area is viewed as a set of *back-office* operations and hidden from customers (Menor, Tatikonda, and Sampson, 2002). The customer-facing service concept prioritizes customer experiences and satisfactory customer-service quality. The service operations/delivery concept (Edvardsson and Olsson, 1996) is concerned with efficiencies and volume. It is suggested that this is where internal newness occurs and the chosen service concept is implemented. Of course, both these concepts must work together as an integrated whole. Efforts to align operational goals with customer preferences are emerging.

There has been limited empirical examination of how organizations design the development process. A notable exception is the early and much-cited work of Shostack (1984) who proposed a way to blueprint, in diagrammatic form, the tangible and intangible elements of the process. In effect, such models determine the prerequisites of the service development because the development is only realized through customer interaction.

The nature and effectiveness of late-stage activities, often termed *implementation*, have been studied from a number of perspectives. Several authors highlight the necessity of customer and staff training in implementing the service innovation. This is perceived in terms of customer use of the service innovation and managing customer expectations. Again, scholars point to the paucity of market testing by service organizations. It appears that the benefits of rapid postlaunch feedback, through ongoing customer interactions in service encounters, can

go some way toward compensating for the lack of formal testing procedures.

Success Factors for Service Innovation Management

Relative to the research on success factors for new "products" (*see* SUCCESS FACTORS FOR NEW-PRODUCT DEVELOPMENT), which is deemed to have reached maturity, research engaged in understanding and articulating the factors that lead to success in service innovation is fragmented. Service innovation research has typically been of a qualitative and descriptive nature. Large-scale empirical studies, with broad samples investigating the antecedents to service innovation performance, have been minimal. A limitation to conducting such types of research lies in the problematic evaluation of service innovation performance or outcomes (*see* INNOVATION METRICS). Managers themselves struggle to develop financial criteria to measure service innovation performance, although it remains the dominant measure in practice. Service managers tend to be less systematic in the way they evaluate innovation than in tangible product scenarios. Further, determining the units of analysis in service innovation measurement is problematic. Innovation can be measured at a project or overall development/portfolio level (Johne and Storey, 1998). Metrics may also be used for the service innovation process or the actual service outcome (or both). Reflecting the nature of services, it is generally viewed that soft customer-based indicators (such as customer satisfaction) and internal measures, such as efficiency of service delivery and effectiveness of front-line staff, are more effective. These metrics are more extensively employed in service organizations.

Notwithstanding such performance-measurement problems, a small number of studies have examined factors correlated with success and failure in new services. Although such studies struggle to indicate generalized success factors, there are a few exceptions. de Brentani (1995), for example, found that success is generally associated with offering services that respond to market needs, reflect the organization's reputation and resources, and are aligned with some form of service-development capability, such as implementation of a formal development process. A limited number of recent large-scale studies support the positive role of a structured formal innovation process. They have found additional evidence for the positive role of technology and the use of interfunctional teams on the effectiveness of service innovation. Research has also proved that differences in the significance of various service innovation success factors depend on contextual factors, such as the level of service innovativeness (van Riel, Lemmink, and Ouwersloot, 2004).

Environment. Apart from such studies reported above, research on factors influencing service innovation success has largely been conducted in a piecemeal fashion, utilizing anecdotal or case-based evidence. Drawing from such studies, we can conclude that considerable attention has been paid to the organizational environment or climate in service innovation. Some studies have pointed to a supportive or positive environment as an important prerequisite to successful management of service innovation. Specifically, the support and commitment of senior management to service innovation, in terms of setting clear goals and providing necessary resources, has been highlighted (Johne and Storey, 1998). Empirical work suggests that senior management can play a critical role in enabling service innovation, such as providing structures that support information sharing or task rotation. Service studies have also pointed to the need for a clear strategic direction or corporate vision for innovation.

Employees. Substantial empirical research lends support to the notion that the service organization's staff contribute to service innovation success, as they are integral to implementation and development. Employees fulfill a critical role in recognizing opportunities for innovation, through their contacts and knowledge of customers and competitors. They embody the new service at the moment of delivery. They can differentiate the service from competitors and help clients make the switching decision (Atuahene-Gima, 1996). However, employees are often reluctant to engage in innovation activities as it increases their workload. Research has explored management techniques to encourage and support staff-based innovation activities.

Attention is paid to the need for training, appropriate reward systems, and motivation techniques (Edvardsson and Olsson, 1996). The role of product champions who push the project through the innovation process has also been found to be important. Recent work has explored the way employees and teams work together and interact in service innovation. Drawing on knowledge and learning paradigms, this emerging school of thought focuses on competencies or routines in information or knowledge sharing among employees (van Riel, Lemmink, and Ouwersloot, 2004; Leiponen, 2006).

External actors. Despite the close interaction between supplier and customer in service delivery, empirical research into the role of customers in the service innovation process is embryonic. Since the 1980s the new-product development literature has suggested that close supplier/customer collaboration can bring success (*see* VOICE OF THE CUSTOMER). Concepts drawn from this work has only, since the late 1990s, been applied or explored in the service innovation domain. Traditionally, the customer is perceived as an inactive component of service innovation. The service organization has been tasked to encompass a customer orientation in the design of the innovation process or operations (Edvardsson and Olsson, 1996). This presupposes that the organization has a comprehensive understanding of its customers' needs. More recently, services researchers have explored an increasingly active role of the customer in service innovation. Research has shown that customers can be involved at different stages of the development process, with varying degrees of intensity, but with greater emphasis on involvement at the early stages of service innovation (Alam and Perry, 2002). Similar studies have unraveled the specific objectives of user involvement in innovation, such as time reduction and speedier service acceptance in the marketplace. Such work has articulated the particular innovation activities that customers perform. Explicit customer co-creation in idea-generating activities propels customers toward determining and, to some extent, designing the value they desire. It has been shown that this can lead to more innovative ideas and service outcomes (Matthing, Sanden, and Edvardsson, 2004).

This greater understanding of the role of customers in the innovation process has fueled further investigation into the way service organizations integrate a broader set of external actors or stakeholders into their innovation process. Innovation in services can be perceived as an outcome of an amalgam of interactions and relationships with numerous actors outside the organization, such as customers, suppliers, distributors, and intermediaries. Managing innovation with a set of external relationships is complex and has started to receive scholarly attention (Syson and Perks, 2004). Organizations are increasingly moving beyond firm boundaries to innovate. Theoretical debate and empirical studies advancing the notion of interorganizational or open innovation (*see* OPEN INNOVATION) abound in the new-product domain. Managing interorganizational service innovation is likely to become a fertile area for research.

Conclusion

To date, most researchers have exploited findings from product innovation to better understand the management of service innovation. While our knowledge of product innovation practices is becoming mature, and a substantive theory is emerging, there is considerable debate around the applicability of such advances to service innovation. There are core intrinsic differences between services and physical products. Scholars have drawn from service management and marketing research streams to articulate these differences, which are centered on the outcome from service innovation and the characteristics of the services themselves. Service innovation researchers are moving beyond replication of product innovation advances and service-specific innovation constructs are emerging. However, rarely do organizations offer pure services or physical products. Increasingly, the offerings of organizations combine both product and service elements, and innovation occurs within both domains. Adoption of traditional "product" or "service" models, or the integration of both, is unlikely to capture or explain the complexities inherent in managing innovation in such contexts. To address these

dichotomies, scholars are beginning to draw on new paradigms to understand and explain the management of service innovation. Although still lagging behind "product" innovation in this respect, the incorporation of new knowledge areas into the discipline will ensure that service innovation management is studied from different perspectives. Such advances will build the foundation for the maturity of the discipline.

Bibliography

Alam, I. and Perry, C. (2002) A customer-oriented new service development process. *The Journal of Services Marketing*, **16** (6), 515–534.

Atuahene-Gima, K. (1996) Differential potency of factors affecting innovation performance in manufacturing and service firms in Australia. *Journal of Product Innovation Management*, **13**, 35–52.

Booz, E., Allen, J., and Hamilton, C. (1984) *New Product Management for the 1980's*, Booz, Allen and Hamilton, New York.

de Brentani, U. (1995) New industrial service development: scenarios for success and failure. *Journal of Business Research*, **32**, 93–103.

Easingwood, C. (1986) New product development for service companies. *Journal of Product Innovation Management*, **3** (4), 264–275.

Edvardsson, B. and Olsson, J. (1996) Key concepts in new service development. *Service Industries Journal*, **16** (2), 140–164.

Fitzsimmons, J.A., and Fitzsimmons, M.J. (2001) *Service Management*, 3rd edn, McGrawHill, New York.

Johne, A. and Storey, C. (1998) New service development: a review of the literature and annotated bibliography. *European Journal of Marketing*, **32** (3/4), 184–251.

Johnson, S.P., Menor, L.J., Roth, A.V., and Chase, R.B. (2000) A critical evaluation of the new service development process: integrating service innovation and service design, in *New Service Development, Creating Memorable Experiences* (eds J.A. Fitzsimmons and M.J. Fitzsimmons), Sage Publications, Thousand Oaks, pp. 1–32.

Leiponen, A. (2006) Managing knowledge for innovation: the case of business-to-business services. *Journal of Product Innovation Management*, **23**, 238–258.

Matthing, J., Sanden, B., and Edvardsson, B. (2004) New service development: learning from and with customers. *International Journal of Service Industry Management*, **15** (5), 479–498.

Menor, L., Tatikonda, M.V., and Sampson, S.E. (2002) New service development: areas for exploitation and exploration. *Journal of Operations Management*, **20**, 135–157.

Page, A.L. and Schirr, G.R. (2008) Growth and development of a body of knowledge: 16 years of new product development research, 1989–2004. *Journal of Product Innovation Management*, **25**, 233–248.

van Riel, A.C.R., Lemmink, J., and Ouwersloot, H. (2004) High-technology service innovation success: a decision-making perspective. *Journal of Product Innovation Management*, **21**, 348–359.

Shostack, G.L. (1984) Designing services that deliver. *Harvard Business Review*, **1**, 133–139.

Syson, F. and Perks, H. (2004) New service development: a network perspective. *Journal of Services Marketing*, **18** (3), 255–266.

organizing for innovation

Richard Blackburn

Introduction

> There are actually two halves to getting innovation to the market. One is having a good idea, which is obviously the easier part. And the second is actually getting that idea through the organization. We need to know how ideas take life, and in such a way that they can be navigated through all of the complexities and competing pressures for resources and focus and points of view that all organizations have (p. 3). Tim Brown, CEO/President of IDEO in Edmonson and Feldman (2005).

> Being able to be empathic, not only to the end user but to the organization you're working with, is critical to this kind of work. The thing that kills most ideas is not that they're not good ideas, it's that they are not good ideas organizationally or culturally (p. 7). Tim Brown, CEO/President of IDEO in Edmonson and Feldman (2006).

IDEO may be as responsible for the phrase "designing for X" as any design consultancy in the world. IDEO's employees have created awareness within many organizations of the fact that products and services can be designed to achieve a variety of outcomes. Thus, in the phrase above "X" could serve as a stand-in for manufacturing, sustainability, safety, aesthetics, quality, cost, and so on. Picking one or more of these terms as the basis for a firm's design philosophy gives those involved with a design project valued guidance as to how they should proceed.

Similarly, the phrase "organizing for X" can also be used by organizations to determine the underlying values guiding how an organization chooses to put itself together. Firms can organize themselves for efficiency, for flexibility, for quality, for speed to market, and so on. For the purpose of this article, replace "X" with the word "Innovation." The article examines the options that organizations have available for putting themselves together in anticipation of providing innovative products or services to their competitive environment in a timely fashion.

Key Definitions

Before beginning this discussion, it would be helpful to provide a couple of key definitions as well as delimit the boundaries of this discussion. In particular, Amabile's (1988) definitional distinctions between CREATIVITY and innovation (*see* INNOVATION TYPOLOGIES) are applied. She defined *creativity* as "the production of novel and useful ideas by an individual or small group of individuals working together" (p. 126). Innovation uses the results of the creative process as its raw material. Thus, from Amabile's perspective, "Organizational innovation is the successful implementation of creative ideas within an organization" (p. 126). For the purpose of what follows, creativity is taken as a given. The issue of interest here is how best to organize to insure the successful implementation of already developed creative ideas.

This article views the organizing process in the same vein as do others in the field, by equating organizing for innovation with how organizations align a variety of important organizational characteristics in service of innovation. This alignment process might also be referred to as *designing the organization* – in this case designing the organization for innovation. A far more expansive view of organizing is taken here than is by some. Organizing is more than simply understanding the organizational structures which might be more or less appropriate for innovation. Organizing, for the purposes of this article, includes consideration of a number of important organizational design characteristics that collectively influence the success an organization has in its innovation process.

Organizations function within a given context and can be described along a number of different characteristics. For example, organizations exist and hope to survive and prosper within a competitive environment. That environment can be described along a number of dimensions that affect managerial perceptions of the levels of certainty and uncertainty about what might happen in this environment in the future. Despite some organizations' best efforts, many aspects of the environment permeate the boundary of the organization affecting directly or indirectly input, throughput, and output

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

subsystems that keep organizations functioning. Organizations also operate within an external cultural milieu. Increased globalization requires that careful attention be paid toward the cultural differences in markets and in operations.

A more important cultural aspect with regard to innovation, however, is the concept of "corporate culture." A corporate culture reflects the underlying values, assumptions, and mores that determine how employees are likely to feel and behave (Daft, 2007). Clearly, cultures can value either innovation or tradition. Innovative cultures are considered in greater detail later in this article.

Given perceptions of the external environment and culture, and limited to some extent by the assumptions and values of the corporate culture, top management teams have a number of other design-related or organizing-related decisions to make. First, they must choose the appropriate goals for the organization. Survival is typically the unstated formal goal, but beyond that goals might be related to becoming the largest, the fastest, the cheapest, and so on. Having identified a goal or set of goals, those at the top now have to determine the strategy that must be used to achieve those goals. For example, profitability as a goal can be achieved via a low margin–high volume strategy or a high margin–low volume strategy. Implementing the chosen strategy is often a challenging task, but in so doing top managers must make additional decisions about the appropriate structure, technology, systems, and so on, best able to effectively implement the strategy in pursuit of key organizational goals. Finally, organizations cannot exist without the ingredient that brings all of these elements to life – their people. Without the employees, the organization is merely an empty shell.

The key to organizational success, of course, is for top management to make the right choices regarding each of these important organizational decisions. The key to successful innovation is making the right choices with regard to accomplishing that particular strategy. And regardless of the individual choices to be made, one important overall guiding principle is that the nature of the decisions made and actions taken with regard to these various aspects must all *fit*. For example, choosing the wrong goals (goals that do not fit), given the environmental and cultural factors, will likely lead to a less-than-successful organization. Should changes be made in goals, for instance, strategies, structures, technologies, systems, and people will likely have to change as well to insure achievement of this revised goal.

When one considers how best to organize for innovation, the puzzle becomes quite complex. To place realistic boundaries around this particular entry, this article focuses on only a subset of the characteristics considered above. It is assumed that someone reading this article has already appropriately examined the competitive environment (*see* CORE COMPETENCIES; COMPETITIVE ADVANTAGE) and made the determination that one of the most important goals for the organization is innovation. The strategy chosen to achieve that goal requires decisions about where to seek/find innovations (internal or external to the organization) (*see* LEAD USERS;OPEN INNOVATION), the timing for innovative products and services (first mover or fast follower, for instance) (*see* FIRST-MOVER ADVANTAGE), and the determination of those participating in the innovation process (internal networks or networks between/among other organizations). Those decisions then necessarily influence decisions about the structure, technology, systems, people, and culture that would best serve to support this innovation strategy in pursuit of the innovation goal. It is assumed that technology, systems, and people are given. Certainly, much could be written about the various technologies and systems that support innovation, but that is beyond the scope of this article. Likewise, the impact of individuals on innovation and the innovation process is substantial, but beyond the focus of this article. The literature about individual creativity is voluminous (*see* CREATIVITY). For all three characteristics, the reasonable assumption made is that the organization will find the technologies, systems, and people necessary to fit with the other aspects in this effort aimed at designing/organizing for innovation. The focus, over the remainder of this discussion, is on two issues: "culturing" for innovation and structuring for innovation.

Innovation Cultures

Daft (2007) defines *culture* as "the set of values, norms, guiding beliefs, and new understandings that is shared by members of an organization and is taught to new members. It represents the unwritten, feeling part of the organization" (p. 361). Culture reflects the observable aspects and behaviors of how an organization functions (the artifacts, symbols, and ceremonies known to employees as well as the nature of personal interactions) as well as the unobservable aspects that reflect the deeper values and assumptions that guide organizational behavior but are quite difficult to identify. In fact, asking employees to discuss these deep-seated values and assumptions is difficult, because it would not be unlike asking a fish to describe water. Its ubiquity makes it nearly impossible to adequately and accurately describe it.

Burns and Stalker argued in 1966 that culture was central to organizing for innovation. They noted that firms need to develop "'codes of conduct,' which would enable people to 'comprehend more eventualities and more information . . . and [in which] the limits of feasible action could be set more widely.'" (Burns and Stalker, 1966, p. 11 quoted in Dougherty, 1999, p. 182) So this concept and how it relates to organizing for innovation is first discussed.

Organizations typically have multiple cultures. However, there is likely an overall set of values and assumptions that guide nearly all employee behaviors through the firm. This overall culture has been greatly influenced both by the societal culture within which the organization was founded as well as the individual values and assumptions of founders and/or key leaders who influence(d) selections of corporate goals, strategies, and employees.

While it is possible to change a corporate culture, it is a very difficult task, given the strength of most cultures. Within this overall corporate culture, one can also find a series of subcultures developed as a function of, say, hierarchical level and/or work unit. Some cultural assumptions guiding behaviors of those at the top of the organization may differ from assumptions guiding behaviors of those at the bottom of the firm. Similarly, cultural assumptions within certain functions, research and development, for example, will likely differ from some of the assumptions found in manufacturing units. For instance, some manufacturing organizations may develop strong or weak quality cultures ("Quality is Job 1") or a health-care organization may develop strong or weak safety cultures. Of particular interest in this discussion is what might an innovation culture look like, and how might one be developed as part of the overall effort to organize for innovation?

There is sufficient evidence in the literature to support the general contention that the nature of an organization's culture, its corporate culture, has a substantive impact on a number of dimensions of organizational performance. Starting with Schein (1992) early consideration of the role of culture in organizations up to a more recent consideration of the relative impact of corporate culture versus country culture (*see* GLOBAL PRODUCT DEVELOPMENT) on RADICAL INNOVATION (Tellis, Prabhu, and Chandy, 2009), the concept of culture has been linked to organizational performance. One need look no further than the differences in cultures that appear to exist between GM and Chrysler and Google, 3M, Apple, Sony, and so on, to appreciate that the nature of the innovation cultures among these organizations has greatly influenced their successes or lack thereof. Even within an industry, innovation culture differences seem to exist. The lay person likely sees some difference in the innovation cultures in Apple versus Microsoft that are reflected in a number of characteristics that distinguish each organization including their marketing, products, and packaging, for instance.

A variety of studies have advanced ideas about what should be the characteristics or dimensions of an innovative culture. The space available in this format limits the number of innovation culture models that can be examined to three. Scott and Bruce (1994) offer a relatively simple model of an innovation culture suggesting that innovation cultures are most often marked by organizational "support for innovation" (openness to change, supportive of new ideas from employees, and tolerant of diversity among employees) as well as "resource support" (the provision of people, time, material, and money in service of innovation).

Table 1 Important characteristics of an innovative culture, by rank.

Factors	*Today*	*In 10 years*
Customer focus	1	1
Teamwork and collaboration with others	2	2
Appropriate resources (time and money)	3	6
Organizational communication	4	3
Ability to select right ideas for research	5	4
Ability to identify creative people	6	5
Freedom to innovate	7	7
Ability to measure results of innovation	8	8
Encouraging both small ideas and big ideas	9	9
Innovation accountability and goals	10	10
Culture of risk tolerance	11	12
Organizational structures	12	11
Diversity	13	13
Balancing incremental improvements and breakthrough discoveries	14	14

Reproduced with permission from AMA/HRI Innovation Survey 2006, p. 18.

At the other extreme, in terms of the number of possible dimensions that might comprise an innovation culture are the results published in 2007 of a survey sponsored by the American Management Association and conducted in 2006 among top executives as to the important dimensions of a culture of innovation (Bear *et al.*, 2006). Table 1 reflects the opinions of over 1300 global respondents to questions regarding the antecedents and consequences of innovation in their organizations now, and their predictions of the relative importance of these same cultural characteristics in 10 years (2016). The authors of this study suggest that there are 14 important characteristics of an innovative culture with customer focus, teamwork/collaboration with others, and the availability of appropriate resources (time/money) identified as the three most important.

Finally, Goffin and Mitchell (2005) propose an innovative culture whose dimensionality falls midway between the extremes considered above. They describe what they label as *the cultural web*. This web, based on the work of Johnson and Scholes (1999), is composed of six aspects of culture supporting an organization's "central paradigm" (pp. 266–268). The central paradigm is a "statement summarizing the main points about how an organization 'thinks' and 'acts.'" Designed to reflect these most basic cultural assumptions are what they describe as the key aspects of culture as captured by the web:

1. *Organizational structure*: the formal organization. The discussion above would argue that rather than being a dimension of an innovation culture, structural choices are instead influenced by the cultural assumptions.
2. *Power structures*: where the power actually resides in an organization. These structures may or may not overlay the formal organizational structure.
3. *Symbols*: advertising, logos, office styles, titles, dress codes, language, and terminology that are unique to an organization.
4. *Stories*: tales that are shared among employees that capture the essence of key events in the firm's history.
5. *Routines and rituals*: the ways in which employees learn to act toward each other and to process work. These often become the taken-for-granted, unquestioned approaches to dealing with various organizational issues.
6. *Control systems*: formal processes, measurement systems, and reward and recognition systems. What these systems assess and

reward reveals what the organization values and on what employee attention should be focused.

For each of these six dimensions, Goffin and Mitchell (2005) offer what they argue are best practices decanted from three major studies of innovation cultures (Jelinek and Schoonhoven, 1990; O'Reilly and Tushman, 1997; Zien and Buckler, 1997). Their suggestions about ways in which the last five of their cultural aspects can support an innovation culture (pp. 274–278) are summarized below. Organizational structure and innovation will be considered separately in the final section of the article.

1. *Power structures*: Substantial power differences tend to reduce innovation and innovative opportunities. A balance of power within a hierarchy and a willingness to cooperate across units contribute to higher levels of innovation. Training and CROSS-FUNCTIONAL TEAM opportunities can encourage this balanced power structure.
2. *Symbols*: All types of communication, internal and external as well as formal and informal, require a focus on innovation: this involves walking the walk and talking the talk; displays of innovations in prominent locations in the organization and on appropriate web pages. The use of unique plaques and certificates to recognize innovators and their innovations should also be publicly available and displayed.
3. *Stories*: "Managers can be transformational leaders (for innovation) if they develop and constantly tell enlightening stories (about innovation) in staff meetings, interviews, and outside speeches" (Goffin and Mitchell, 2005, p. 275).
4. *Routines and rituals*: Programs that encourage the development of new ideas for both product and process improvements may be as structured as internal venture funding of new ideas, encouraging the use of web sites that allow employees to vote on ideas posed by other employees, or allowing new employees to provide formal feedback on what they feel should be done differently. On the other side of the coin, tolerating mistakes and/or establishing a mind-set that failure may be a better way of learning than success are two ways of encouraging the risk taking associated with successful innovation. One of the key cultural assumptions of IDEO, the world-renowned design consultancy, is described by its cofounder Tom Kelly: "Failure is part of the culture. We call it enlightened trial and error" (in Thomke and Nimgade, 2000, rev. 2007, p. 4).
5. *Control systems*: Because what gets measured and rewarded is what gets done, the design of metrics and their link to rewards for innovation must be carefully considered. If innovation is desired, then evaluation and reward systems have to measure and reinforce innovative behavior (*see* INNOVATION METRICS).

Structuring for Innovation

This final section examines what Goffin and Mitchell (2005) consider as part of corporate culture, but which, it is felt, is a derivative of that culture, namely, the organization's structure. As has been the case with innovation and culture, definitions of structure are numerous. For our purposes, we discuss the formal structure of the organization – how the organization puts its various work units together to accomplish its goals.

There are a number of tensions that those at the top of organizations have to manage as they build their firms. For instance, there is the tension between differentiation (taking complex tasks apart so that they can be done more quickly and efficiently) and integration (putting the results of those differentiated tasks back together into a useful good or service). There is the tension between efficiency (producing at the lowest possible cost) and flexibility/adaptability (producing goods and services to meet specific customer needs).

Dougherty (1999) describes four different tensions that organizations confront when they attempt to organize for innovation. The first is the external–internal tension. This tension manifests in two ways. It can refer to the extent to which the market makes certain innovation demands on the organization, only some of

which the internal organization elements are able to provide. This tension can also relate to how organizations manage the extent to which innovation is sought from sources outside the firm or the extent to which firms outsource the development of innovative ideas to other firms (*see* LEAD USERS; VOICE OF THE CUSTOMER; FRONT END OF INNOVATION). How to draw organizational boundaries and how to manage across those boundaries contributes to a variety of possible structures.

Second, there may be a tension between the old and the new. "(T)he complexity of work . . . reinforces a tendency to separate and compartmentalize work, locates power in the established business, not new ones, and hinders integrated problem-solving" (Dougherty, 1999, p. 178). This reflects the tension between exploitation (continuing to reap marginal revenues from minor improvements in the current line-up of goods and services) and exploration (the capability to develop new and different goods and services) (March, 1991).

A third tension may arise between what Dougherty calls *strategic determination* and *strategic emergence*. This refers to the innovation required by top-down strategic decisions versus innovations that are allowed to bubble up from the bottom of the firm. The tension here is a lack of really new ideas if top-down innovation only supports changes at the margin versus the possible inability of most bottom-up innovations to build on previous work and/or to fit with the organization's strategy.

The final tension is the one experienced by those who struggle between the freedom to be innovative and the responsibility to work toward common organizational goals. Larger firms tend to emphasize responsibility over freedom – doing things right; not necessarily doing the right things. Dougherty asks, "How can large, complex organizations solve the problems of normal functioning . . . and still embody the tensions which power innovation" (Dougherty, 1999, p. 181).

As organizations attempt to manage these tensions and organize themselves for innovation, they have a number of structural options available. If efficiency and exploitation are important, then the costs savings associated with a functional structure suggest that functional innovation teams may be beneficial. These include teams of employees located within the same functional unit. This approach frequently limits the innovative ideas to those affecting a single function, with the teams often providing incremental innovations at best to the firm.

Typically, however, many (most) innovations (most new-product or service innovations) are too complex to be isolated within a single function. In this case, the use of CROSS-FUNCTIONAL TEAM may be the appropriate structural solution. These teams may be colocated or, given the precipitous decline in the costs of communication and computer technology, operate virtually where key functional team members operate at a distance from each other (*see* VIRTUAL TEAMS). These cross-functional teams usually operate within an existing organizational structure with team members still responsible to their functional managers, but with only "dotted line" responsibility to the leadership of the cross-functional team. One difficulty that often besets cross-functional teams is that the team manager (if one exists) does not have direct authority over team members, meaning that his/her power, influence, and access to necessary resources may be limited.

One structural solution to this problem is in the formation of the heavyweight cross-functional team. Under this structure, the team leader has direct authority over team members, while functional managers now have "dotted line" responsibility for their respective team members. The team leader's position in this structure is at the same hierarchical level as that of the functional managers supplying team members, and he/she can exercise the necessary power and influence over team members as well as be better positioned to acquire the necessary resources.

Yet another structural option that provides even greater autonomy to those charged with developing innovative products or services is the autonomous team. These teams have strong leaders, and they are structured so as to be independent of the broader organizational hierarchy and culture. Lockheed's attempt to design a new jet fighter in the early 1940s is probably the first formal use of what they dubbed

Table 2 The scope of the ambidextrous organization.

Alignment of	*Exploitative Business*	*Exploratory Business*
Strategic intent	Cost, profit	Innovation, growth
Critical tasks	Operations, efficiency, incremental innovation	Adaptability, new products, breakthrough innovation
Competencies	Operational	Entrepreneurial
Structure	Formal, mechanistic	Adaptive, loose
Controls, rewards	Margins, productivity	Milestones, growth
Culture	Efficiency, low risk, quality, customers	Risk-taking, speed, flexibility, experimentation
Leadership role	Authoritative, top down	Visionary, involved

Reproduced with permission from O'Reilly, C. and Tushman, M. (2004) The ambidextrous organization. Harvard Business Review, April 7.

the "*skunkworks*." It is also the case that any of these team-based structures can operate with colocated team members but can also be structured virtually with team members located at a distance from each other.

These structural options above seem for the most part to be "grafted" on to the existing organization structure. While this is not necessarily a problem, particularly if other key organizational characteristics (importantly, corporate culture) are supportive of the innovation goals and strategies, more recently, some have suggested that organizations consider more substantive structural changes that produce what is described as the *ambidextrous organization* (Daft, 2007; O'Reilly and Tushman, 2004).

One of the tensions or trade-offs that organizations confront in an environment calling for more effective innovation is between the flexibility of the organic design and the structure of the mechanistic design. As Daft notes, "The challenge for managers is to create both organic and mechanistic conditions within the organization to achieve both innovation and efficiency" (Daft, 2007, p. 407). The issue is that while organic organizations are better at getting employees to generate innovative ideas, they are not the best means for actually implementing ideas. On the other hand, mechanistic designs may be required to actually implement innovative ideas once they have reached that stage of the development process. As O'Reilly and Tushman (2008) note, "... the routines, processes, and skills required for exploitation are fundamentally different than those required for exploration ..." (p. 188).

Enter the ambidextrous organization. This type of organization is designed to behave organically when exploring for new ideas and mechanistically when determining how to exploit these innovations. This design or structure can be applied within a unit or between units within the larger organization. In their research, O'Reilly and Tushman (2004) found that organizations that have been successful in both exploiting the success of their present product/service line(s) and exploring future innovations structurally separate the exploratory units from the exploitative ones. But "at the same time, they maintain tight links across units at the senior level" (O'Reilly and Tushman, 2004, p. 2). Table 2 provides information as to the design differences that arise in the units that might comprise an ambidextrous organization.

Each of these structural options comes with its own set of advantages and disadvantages, and the choice of the appropriate structure depends on how well a particular structural approach fits with the other elements of the organization's design. For instance, Daft (2007) suggests that the culture of an organization should reinforce the strategy and structure of an organization (and vice versa). The importance of the appropriate fit among the various organizational characteristics identified above should not be underestimated.

Conclusion

In summary, Kanter (1988) asserts that by its nature the innovation process is uncertain, knowledge intensive, controversial, and cross-functional. If these characteristics of innovation are still accurate, then her remarks made more than 20 years ago are still true today as we think about the best ways to organize for innovation.

> It (innovation) is most likely to flourish where conditions allow flexibility, quick action and intensive care, coalition formation, and connectedness. It is most likely to grow in organization that integrated structures and cultures emphasizing diversity, multiple structural linkages both inside and outside the organization, intersecting territories, collective pride and faith in people's talents, collaboration and teamwork (p. 172).

Similarly, the conclusions drawn by the authors of the AMA/HRI Innovation Survey are also compelling:

> The future challenge for companies will be to develop an 'agile mindset' that allows them to quickly respond to changes in the marketplace, new technologies, geopolitics and other factors. The standard organizational pyramid with the vertical hierarchy of boxes, while not disappearing entirely, is likely to lead to a more over-lapping and highly linked set of satellites connected by information technology. A growing number of organizations will be characterized by an integrated and dispersed set of mobile, multi-functional expert teams rather than by separate functions and distinct regional offices. The key will be to quickly focus and organize resources to support strategic initiatives, including innovation goals. (Bear *et al*., 2006, p. 45).

Bibliography

Amabile, T. (1988) A model of creativity and innovation in organizations, in *Research in Organizational Behavior*, vol. **10** (eds B. Staw, and L. Cummings), JAI Press, Greenwich, pp. 123–167.

Bear, D., Chaves, W., Conte, S. *et al*. (2006) *The Quest for Innovation: A Global Study of Innovation Management 2006–2016*, American Management Association, New York.

Burns, T. and Stalker, G. (1966) *The Management of Innovation*, 2nd edn, Tavistock Institute, London.

Daft, R. (2007) *Organization Theory and Design*, 9th edn, South-Western Cengage Learning, Mason, pp. 398–440.

Dougherty, D. (1999) Organizing for innovation, in *Managing Organizations: Current Issues* (eds S. Clegg, C. Hardy, and W. Nord), Sage Publications, Inc., Thousand Oaks, pp. 174–189.

Edmonson, A. and Feldman, L. (2005), rev. 2006 *Phase Zero: Introducing New Services at IDEO (A)*, Case Study #9-605-069, Harvard Business School Publishing, Cambridge.

Edmonson, A. and Feldman, L. (2006), rev. 2007 *Phase Zero: Introducing New Services at IDEO (B)*, Case Study #9-605-069, Harvard Business School Publishing, Cambridge.

Goffin, K. and Mitchell, R. (2005) *Innovation Management: Strategy and Implementation Using the Pentathlon Framework*, Palgrave Macmillan, New York, 275.

Jelinek, M. and Schoonhoven, C. (1990) *The Innovation Marathon: Lessons from High Technology Firms*, Basil Blackwell, New York.

Johnson, G. and Scholes, E. (1999) *Exploring Corporate Strategy*, 5th edn, Pearson Education Ltd, Edinburgh.

Kanter, R. (1988) When a thousand flowers bloom: structural, collective, and social conditions for innovation in organizations, in *Research in Organizational Behavior*, vol. **10** (eds B. Staw, and L. Cummings), JAI Press, Greenwich, pp. 169–211.

March, J. (1991) Exploration and exploitation in organizational learning. *Organization Science*, **2**, 71–87.

O'Reilly, C. and Tushman, M. (1997) Using culture for strategic advantage: promoting innovation through social control, in *Managing Strategic Innovation and Change: A Collection of Readings* (eds M. Tushman, and P. Anderson), Oxford University Press, New York, pp. 200–216.

O'Reilly, C. and Tushman, M. (2004) The ambidextrous organization. *Harvard Business Review*, **82** (4), 1–8.

O'Reilly, C. and Tushman, M. (2008) Ambidexterity as a dynamic capability: resolving the innovator's dilemma, in *Research in Organizational Behavior* (eds A. Brief, and B. Staw), Elsevier Ltd, New York, pp. 185–206.

Schein, E. (1992) *Organization Culture and Leadership*, 2nd edn, Jossey-Bass, San Francisco.

Scott, S. and Bruce, R. (1994) Determinants of innovative behavior: a path-model of individual innovation in the work place. *Academy of Management Journal*, **37**, 580–607.

Tellis, G., Prabhu, J., and Chandy, R. (2009) Radical innovation across nations: the preeminence of corporate culture. *Journal of Marketing*, **73**, 3–23.

Thomke, S. and Nimgade, A. (2000), rev. 2007 *IDEO Product Development*, Case Study #9-600-143, Harvard Business School Publishing, Cambridge.

Zien, K.A. and Buckler, S.A. (1997) Dreams to market: crafting a culture of innovation. *Journal of Product Innovation Management*, **14**, 247–287.

lead users

Nikolaus Franke

Who are Lead Users?

Lead users represent valuable external sources for companies seeking ideas and concepts for RADICAL INNOVATION, be they products or services. The need to systematically transcend the company's internal resources when innovating, particularly at what is called the FRONT END OF INNOVATION of the innovation process, is highlighted in the literature on OPEN INNOVATION.

Lead users are defined as members of a user population who display two key characteristics: first, they anticipate relatively high benefits from obtaining a solution to their needs – and may innovate as a result. Second, they are at the leading edge of important trends in a given marketplace – and so are currently experiencing needs that will later be experienced by many users in that marketplace (von Hippel, 1986).

In a number of studies from the late 1970s and 1980s onward, Eric von Hippel of Massachusetts Institute of Technology (MIT) and several of his colleagues have observed that in very different industries – ranging from high-tech areas such as scientific instruments or thermoplastics to consumer markets such as outdoor equipment or skateboards – a huge percentage of the most important innovations were originally developed by the product users, not by the producing firms (see von Hippel, 1988; von Hippel, 2005 for overviews). In this context, the term *user* refers to the functional role of the institution and means that with respect to the product in question, the institution expects to derive benefits from its own product *use*, not from *selling* the product. Therefore, "users" may be individual end users (e.g., consumers in the food market) or firms (e.g., a high-tech manufacturer who uses a specific machine in its internal production process).

The finding that users can be very active in innovation seemed to contradict canonical market research experience from "VOICE OF THE CUSTOMER" techniques, which holds that customers are at best capable of articulating unsatisfied present needs but are hardly able to provide information about future needs or even to provide ideas, concepts, and solutions to match those needs. This puzzle was resolved by the introduction of the "lead-user" concept. Although it may be true that many customers are unable to provide active input in new-product development tasks, there is a specific subgroup of users – known as *lead users* – who are indeed creative and innovative. Lead users are able to provide direct input in new-product development tasks and have often developed prototypes (*see* PROTOTYPE) of new-product solutions for themselves (personally or for the company they work in) or for their communities (von Hippel, 1986; Franke and Shah, 2003; Lüthje, 2004; Lettl, Hienerth, and Gemünden, 2008). On this basis, Urban and von Hippel (1988) proposed the "lead-user method," which allows companies to benefit from the creative potential of lead users.

Theoretical Foundations of the Lead-user Construct

The original theoretical thinking that led to the definition of "lead users" as having (i) high expected benefits from an innovation and (ii) a position ahead of an important market trend was built on findings from two different streams of literature (von Hippel, 1986; 2005).

The "high expected benefits" component of the lead-user definition was derived from research on the economics of innovation. Studies of industrial product and process innovations have shown that the greater the benefit an entity expects to obtain from a required innovation, the greater that entity's investment in obtaining a solution will be. The benefits a user expects can be higher than those expected by a manufacturer, for example, if the market is new and uncertain, if customer preferences are heterogeneous and change quickly in the market, or if the costs of innovation are lower for users than for manufacturers owing to the "stickiness" of preference information. Naturally, the benefits a user expects will vary depending on situational variables. Component 1 of the lead-user definition was therefore intended to serve as an indicator of innovation likelihood.

The second component of the lead-user definition, namely, being "ahead of an important marketplace trend," was included because

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

of its expected impact on the commercial attractiveness of innovations developed by users residing at that location in a marketplace (von Hippel, 1986). Studies on the diffusion of innovations regularly show that some adopt innovations before others. Furthermore, classic research on problem solving reveals that subjects are heavily constrained by their real-world experience through an effect known as *functional fixedness*. For example, those who use an object, or see it used in a familiar way, find it difficult to conceive of novel uses. Taken in combination, these findings led to the hypothesis that users at the leading edge would be best positioned to understand what many others will need later. After all, their present-day reality represents aspects of the future from the viewpoint of those with mainstream market needs. Component 2 of the lead-user definition therefore indicates the commercial attractiveness of an innovation created by such a user.

Note that these two components of the lead-user definition are conceptually independent. They stem from different areas of literature, and they serve different functions in lead-user theory. Although they may be correlated in some cases and to some degree, especially as a position ahead of the trend may well be accompanied by a high need for innovative solutions, this is not necessarily always the case. Therefore, the lead-user construct can be described as consisting of two (formative) dimensions.

These theoretical considerations were tested and confirmed in a study of 456 kite surfers and their user innovations (Franke, von Hippel, and Schreier, 2006). Figure 1 shows the main findings in graphic form: first of all, we see

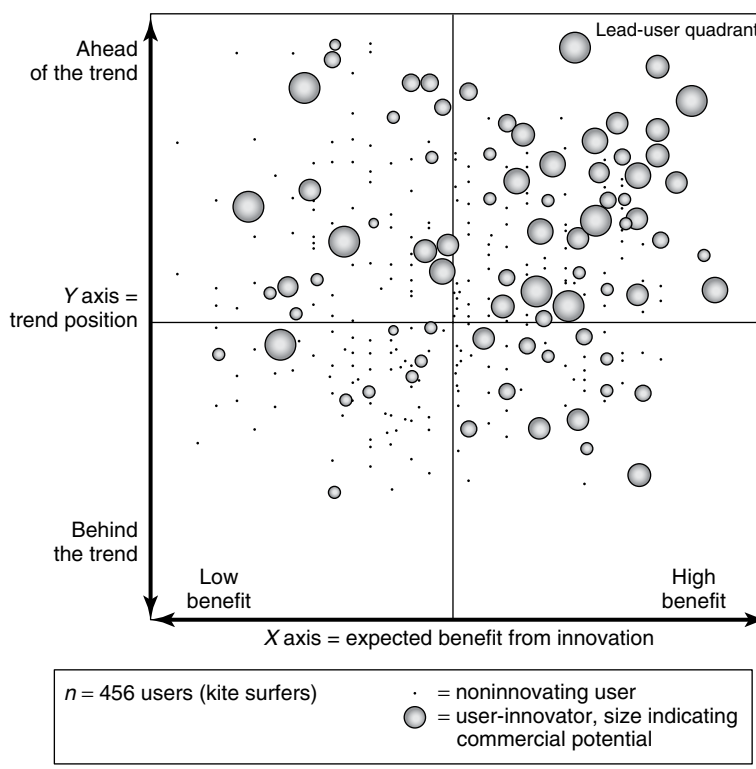


Figure 1 Innovation likelihood and innovation attractiveness as a function of the two lead-user dimensions (adapted from Franke, von Hippel, and Schreier, 2006).

that both components are indeed relatively independent. Users (represented by small dots or bubbles) are broadly distributed, and a considerable number of users are far ahead of the trend but would hardly expect any benefit from innovating. At the same time, many users would derive high benefits from an innovation but are not at all ahead of the trend. Second, moving from left to right (i.e., from low to high benefit), we can see that the proportion of innovators (represented by the gray bubbles) rises relative to the number of noninnovating users (represented by the small dots). Third, moving upwards (i.e., from a position behind the trend to a position ahead of the trend), we can see that the attractiveness of innovations (represented by the size of the gray bubbles) increases. Hence, both the proportion of users with innovative ideas and the commercial attractiveness of the innovations they develop are highest in the lead-user quadrant (top right) of Figure 1.

The graph also illustrates what Morrison, Roberts, and Midgley (2004) discovered, namely, that the lead-user construct is distributed over a continuum: users are not simply "lead users" or "non-lead users" (as some earlier literature appears to suggest); instead, they may have *more or fewer* lead-user characteristics.

Recent studies have found that an individual's "lead userness" with respect to a specific market is correlated with other characteristics, such as innovativeness, adoption behavior, or opinion leadership (Schreier and Prügl, 2008). This suggests that lead users may very well constitute valuable resources in other phases of the innovation process, such as NEW-PRODUCT FORECASTING, PRODUCT TESTING and CONCEPT TESTING, PRODUCT DESIGN, and the DIFFUSION OF INNOVATION.

If lead-user innovators are individuals, the question arises as to how they tackle the often complex task of product development. An individual may well develop an idea, but developing the idea into a functioning prototype often requires diverse and specific knowledge, which a lone individual is unlikely to possess. As a result, lead users often organize into communities in order to complement their capabilities, both in off-line communities (Franke and Shah, 2003) and in online communities such as open-source networks (*see* OPEN SOURCE), or other forms of virtual institutions. This makes lead users easier to identify for firms seeking to benefit from their creative potential.

The Lead-user Method

The lead-user method proposed by Urban and von Hippel (1988) is a managerial heuristic that enables companies to search for commercially attractive user innovations and identify new business opportunities systematically (*see* OPPORTUNITY IDENTIFICATION). Usually, this method is described as consisting of four phases (Figure 2): (i) the start phase, (ii) the identification of major needs and trends, (iii) the identification of users leading those trends, and (iv) the lead-user workshop, in which concepts are designed (von Hippel, Thomke, and Sonnack, 1999; Lüthje and Herstatt, 2004).

The start phase involves defining objectives (e.g., "finding an innovative solution to problem X" or "identifying an innovative product concept in market Y") and setting up a CROSS-FUNCTIONAL TEAM.

In the second phase, the three to five most important trends are selected. This selection is usually based on interviews with experts, information from online forums, and literature research. Their function in the process is to narrow the problem and to allow a systematic search for lead users. The trends are those

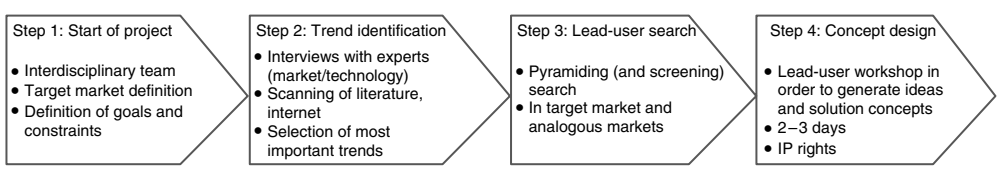


Figure 2 The lead-user method.

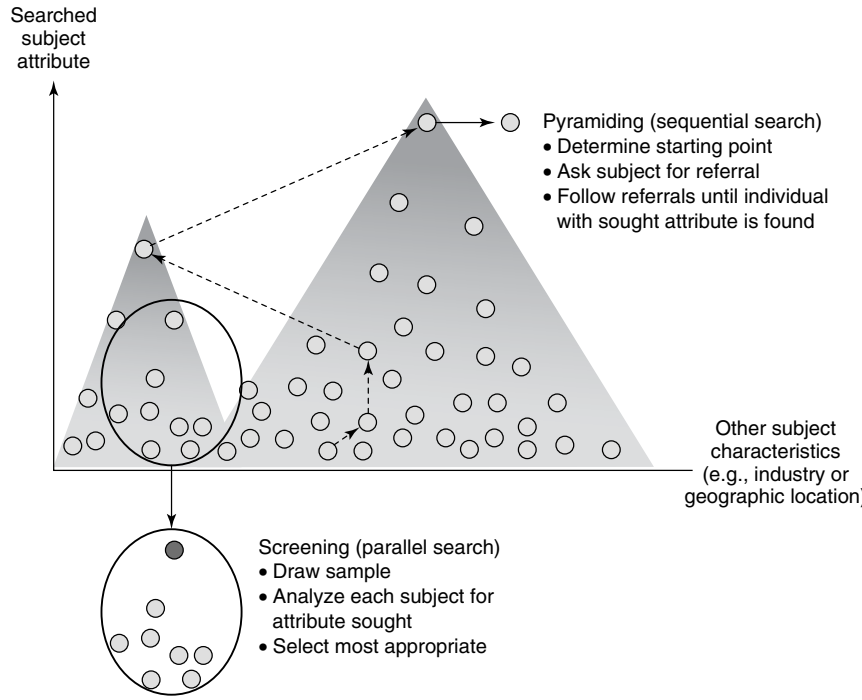


Figure 3 Search strategies: screening and pyramiding (adapted from von Hippel, Franke, and Prügl, 2009).

dimensions in which lead users are far ahead of the mass market.

The third phase involves searching for lead users. Earlier studies usually employed a mass screening approach in which a large sample of users (typically from customer databases) was systematically filtered to identify those users who score highest in both lead-user dimensions. More recently, lead-user studies have increasingly turned to the pyramiding method for the purpose of lead-user identification (von Hippel, Thomke, and Sonnack, 1999). In the latter approach, researchers start with a few users and ask them who has especially high needs and is leading the trend. Those users are then contacted and asked the same questions, and the process continues until a sufficient level of "lead userness" is achieved (which is usually the case after two or three steps). Recently, experiments have demonstrated the superior efficiency of the pyramiding search strategy compared to screening as shown in Figure 3 (von Hippel, Franke, and Prügl, 2009).

Another advantage of the pyramiding approach is the possibility of identifying individuals outside a predefined population or sample. Particularly analogous markets, that is, markets that are different from the target market but characterized by the same trends, are valuable sources in the search for lead users. Consider the example of a lead-user study that aims to find methods of preventing infections in clinical surgery. For this purpose, one important trend would be "methods for increased air purity." Outside of leading hospitals, experts from the analogous field of chip production or CD production may also be able to provide valuable creative input. Two arguments can be brought forward as to why it might make sense to ask such people: first, they might possess solution-related knowledge that is worth transferring from the analogous field to the target field, and second, they are less likely to be blocked by existing solutions in the target field.

In the fourth and final phase, the lead users identified are invited to a two- or three-day

workshop in which company members from different areas (e.g., RESEARCH & DEVELOPMENT, production, and marketing) also participate. At these workshops, classic techniques such as BRAINSTORMING, group discussions, and others are used to capitalize on the CREATIVITY of the participants. It is important for the company to address the issue of INTELLECTUAL PROPERTY RIGHTSintellectual property rights prior to the workshop and to ensure that the ideas and concepts generated can be commercialized without the risk of legal infringements. In many cases, this is less problematic than one would suppose. Harhoff, Henkel, and von Hippel (2003) have shown that, in many cases, it is economically profitable for a user to reveal his/her innovations freely (e.g., because he or she will profit from the use of the resulting product).

Two studies have compared the commercial success of the lead-user method with the output of more traditional methods of idea generation (such as focus groups in quantitative terms and found that the ideas generated by a process using input from lead users exhibit a far higher level of commercial attractiveness (Urban and von Hippel, 1988; Lilien *et al.*, 2002). In the latter study, the authors compared 47 real funded projects at 3M. They found that the projected commercial success of projects based on the lead-user method ($146 million) was, on an average, 8 times higher than that of projects based on traditional methods ($18 million). The projects also differed significantly in terms of novelty, originality, strategic importance, and radicalness. This underscores the potentially high value of lead users in marketing and new-product development.

Bibliography

Franke, N. and Shah, S. (2003) How communities support innovative activities: an exploration of assistance and sharing among end-users. *Research Policy*, **32**, 157–178.

Franke, N., von Hippel, E., and Schreier, M. (2006) Finding commercially attractive user innovations: a test of lead user theory. *Journal of Product Innovation Management*, **23**, 301–315.

Harhoff, D., Henkel, J., and von Hippel, E. (2003) Profiting from voluntary information spillovers: how users benefit by freely revealing their innovations. *Research Policy*, **32**, 1753–1769.

von Hippel, E. (1986) Lead users: a source of novel product concepts. *Management Science*, **32**, 791–806.

von Hippel, E. (1988) *The Sources of Innovation*, Oxford University Press, New York.

von Hippel, E. (2005) *Democratizing Innovation*, MIT Press, Cambridge.

von Hippel, E., Franke, N., and Prügl, R. (2009) "Pyramiding": Efficient Identification of Rare Subjects, Research Policy, forthcoming.

von Hippel, E., Thomke, S., and Sonnack, M. (1999) Creating breakthroughs at 3M. *Harvard Business Review*, **77**, 47–57.

Lettl, C., Hienerth, C., and Gemünden, H.G. (2008) Exploring how lead users develop radical innovation: opportunity recognition and exploitation in the field of medical equipment technology. *IEEE Transactions on Engineering Management*, **55**, 219–233.

Lilien, G., Morrison, P.D., Searls, K. *et al.* (2002) Performance assessment of the lead user generation process for new product development. *Management Science*, **48**, 1042–1059.

Lüthje, C. (2004) Characteristics of innovating users in a consumer goods field. An empirical study of sport-related product consumers. *Technovation*, **24**, 683–695.

Lüthje, C. and Herstatt, C. (2004) The lead user method: an outline of empirical findings and issues for future research. *R&D Management*, **34**, 553–568.

Morrison, P.D., Roberts, J.H., and Midgley, D.F. (2004) The nature of lead users and measurement of leading edge status. *Research Policy*, **33**, 351–362.

Schreier, M. and Prügl, R. (2008) Extending lead user theory: antecedents and consequences of consumers' lead userness. *Journal of Product Innovation Management*, **25**, 331–346.

Urban, G. and von Hippel, E. (1988) Lead user analyses for the development of new industrial products. *Management Science*, **35**, 569–582.

intellectual property rights

Ludmila Striukova

Definition

Intellectual property (IP) refers to *intangible assets* related to knowledge and ideas. Protection of IP differs from protection of tangible assets, as intellectual assets are generally costly to create, but easy to copy. IP is protected by registered rights, such as design rights, trademark and patents, and unregistered rights, such as copyright and trade secrets.

The main rationale behind intellectual property rights (IPRs) is to protect the inventor and to ensure that the inventor can appropriate sufficient returns from innovation, and to encourage further innovation. Indeed, strong IP protection can be beneficial to both small (Grindley and Teece, 1997; Gans and Stern, 2003; Rivette and Kline, 2000) and large (Kalamas *et al.*, 2002) companies. On the other hand, there is evidence that a too rigid IP regime can hinder innovation (Merges and Nelson, 1990; Scotchmer, 1991).

Copyright

Copyright protects original literary, dramatic, musical, and artistic work. However, it protects the expression of ideas, rather than the ideas themselves. The original copyright law in England and Scotland (1709) gave copyright for 14 years plus additional 14 years if the author was still alive. The duration of copyright, however, has since been increased. Nowadays, in most countries, copyright lasts for the life of the author plus 70 years. Furthermore, recordings are protected for 50 years in Europe and 95 years in the United States. Main contemporary issues related to copyright are the digitalization of data and software-related issues. and software-related issues are the main contemporary issues involving copyright.

While the protection provided by copyright has been considered for centuries as rather adequate, its effectiveness has been recently challenged by the advent of digital technology. Traditionally, copyright was infringed only by a limited number of individuals or companies, which is one of the reasons for its success. However, the fact that digital goods can be reproduced at nearly no cost without losing quality has enabled almost anybody to become an infringer. Because of that, the protection of copyright has become an issue for music and film industries (among others) who claim that they lose millions of dollars per year because of the illegal distribution of digital goods and are in a constant struggle to stop individuals from sharing copyrighted content among themselves. Napster, for instance, one of the first popular file-sharing services, was shut down in 2001 when it was found liable of infringement, not because Napster itself infringed any copyright, but because it enabled copyright infringement.

Similarly, in 2005, the Authors Guild, the Association of American Publishers and individual authors and publishers sued Google, which scanned books and allowed its users to search through them. The lawsuit was settled in October 2008 when Google agreed to pay compensation to the copyright holders whose books had already been scanned, to cover the cost of creation and maintenance of a register with information about books and their usage. Also, copyright holders were given control over how much of the content is included and they now receive a percentage of the revenues generated by this activity.

Though the outcome of the suit provides copyright holders with a remuneration, it also undermines one of the most important goals of Google Books, which was to provide access to rare or out-of-print books or to provide access to those who live far from bookstores or public libraries. This may happen if the copyright holders choose not to include their books or if the books are "orphan works," that is, when copyright owners cannot be identified. Furthermore, the settlement does not allow short snippets of text to be included in Google search, despite this falling under "fair-use" category of copyright law.

Because of the steep increase in the number of copyright infringers and the quasi impossibility of initiating litigation against all of them, copyright holders have changed strategy and have attempted in advance to prevent infringement from taking place. Digital right management (DRM) is a set of technologies that aims to

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

prevent consumers from illegal copying and sharing. The general idea of DRM is to distribute encrypted media files that need to be decrypted before use. Such decryption requires connecting to a central server that holds a database of legitimate users and of rights each user has in regard to the usage of a particular digital product. While DRM was originally designed as fit to curb piracy, there are numerous ways to circumvent existing DRM systems (Rayna and Striukova, 2008a).

In addition, largescale adoption of DRM technologies needed for it to impact on piracy has been hindered by concerns such as the breach of privacy caused by personal data necessarily collected by the central servers as well as the lack of flexibility introduced by these technologies. Technologies such as DRM make it possible to collect personal data on regular basis even without customers' knowledge and to be used for price discrimination. Rayna and Striukova (2008b) propose two solutions to the increasing problem of loss of piracy: keep the current DRM technology, but to reward customers for disclosing their personal information, or to establish anonymous and more flexible DRM systems.

Another current issue associated with copyright is related to the protection of software. The source code of software is protected by copyright, however, there have been cases when the structure of the program and its "look and feel" have been copied and while, the source code was not copied and differed between the two pieces of software, this has nonetheless be ruled out as a copyright infringement.

Trade Secret

A trade secret is a piece of information that provides an organization with a competitive advantage over other companies; for example, formulae, plans, and manufacturing processes. It is indefinite, however, as with any secret, this advantage is lost if someone (presumably a competitor) uncovers it. A company's projected sales figures may also constitute a trade secret because that information is valuable to a competitor.

A good example is the formula for Coca-Cola, which has been kept a trade secret for more than 120 years. The strong protection against competitors provided by the trade secret led to Coca-Cola not being sold in India until 1991, because Indian laws required, before this date, that the formula had to be revealed prior to commercialization. Revealing trade secrets can be made illegal by making employees sign nondisclosure agreements. In 2007, three employees of Coca-Cola received prison sentences for trying to sell, before its release, samples of a new flavor of Coca-Cola to PepsiCo for $1.5 million.

Design Protection

Design rights protect the appearance of the whole or part of a product, in particular contours, shape, and texture. Design protection can be registered and, subsequently, unregistered. In Europe, registration of industrial design can last up to 25 years, subject to renewal every 5 years. In Japan, registration of a design grants a protection for 15 years. In the United States, design is protected by utility patents. In addition, in the United Kingdom, unregistered design is protected automatically for 15 years. It can be noted that two-dimensional design (e.g., a drawing or a logo) also can be a subject to copyright protection. However, copyright protects only the two-dimensional object (e.g., a sketch of a dress) and cannot be applied when the project becomes three-dimensional (a dress itself). Design protection is important for many industries; however, some industries, like the fashion industry, do not impose strict enforcement. First of all, the life of a fashion design is so short that it does not make sense for companies to apply for design protection and to subsequently enforce it; furthermore, copying in fashion industry speeds diffusion and makes the obsolescence of fashion designs more rapid, resulting in new models and designs (Raustiala and Sprigman, 2006).

Trademark

Originally, the purpose of a trademark was to certify the origin of goods. However, nowadays, its role has changed, since trademarks are often licensed out and one company can hold multiple trademarks. In practice, a trademark is used by companies to attract consumers when the brand

name protected by the trademark has earned a positive reputation. To be of value, however, a trademark should be known and appreciated by a large number of people; maintaining this may therefore be quite costly (Smith, 1997).

A trademark is initially granted for 10 years and can then be renewed indefinitely as long as it is in use. If a trademark becomes generic (i.e., the name has become a generic designation of a particular kind of product in the popular language, such as hoover, elevator) it may be revoked. This may also happen if a trademark is no longer in use.

The breadth of trademark has been enlarged by the advent of the Internet. Indeed, domain names are now considered as an extension of the trademark and, therefore, priority is given to the trademark holder when an application is made for a domain name.

Patent

A patent grants an inventor limited monopoly in exchange for a disclosure of information about their invention. There are four broad theories about the principal purposes of patents (Mazzoleni and Nelson, 1998): patents motivate useful invention and without a strong patent system there will be no invention, facilitate wide knowledge about the use of inventions by inducing inventors to disclose their inventions when otherwise they would rely on trade secrets, induce the investment needed to develop and commercialize inventions and enable the exploration of broad prospects. Granting broad patents on early stage inventions enables inventors to carry out subsequent research and development within the area of the patent claim and be protected by the same patent.

Patents last for 20 years from the date of filing. To be granted a patent, an invention has to be new, nonobvious (novel), and capable of industrial application. The nonobvious clause was only introduced in the United Kingdom in 1902 and in the United States in 1966 (Graham vs. John Deere).

In the United States, the rights are given to those who were first to create an invention and not to those who were first to file for a patent. Also, in the United States, the inventors have one year after they commercialize their invention to file for a patent. Patents are a territorial right and have to be filed for, in different countries. In the European Union, the patent system has moved toward a greater harmonization and the European Patent Office (EPO) was opened in 1978 to centralize and ease patent applications. However, European patents still do not exist and applicants still need to apply (and pay) for different countries. Yet, the creation of the EPO has facilitated the application process by establishing a one-stop central application service in Europe. It is to be noted that the EPO is only in charge of granting patents, which then have to be enforced separately in the different countries. Also, sometimes, a patent protecting one invention can be valid in one country, but judged to be invalid in another country. Until recently, patents had to be translated in all European languages, however, since May 2008, most of the European countries relinquished the requirement that patent has to be translated into their national language, saving patent applicants up to 40% in fees.

The domain of application of patents has evolved over time. In the United States, it became possible to patent software in the 1980s, and business methods in 1998. In the European Union, "pure" business methods are not patentable, however, if a method is technical it can be related to doing business and still be patentable.

Patent Management

The changing role of patents made companies take a more strategic approach to patent management (*see also* LAUNCH STRATEGIES; GROWTH STRATEGIES; MANAGING MATURE PRODUCTS). While patents have existed now for a long time, the value that resides in them has long been overlooked by firms and it is only recently that companies started to create special departments to acquire and manage IP, aiming at maximizing value through IPRs. According to Reitzig (2003), factors affecting the value of a patent include its lifetime, its breadth, and family size, disclosure, the difficulty of inventing around, legal argument, and backward and forward citations. The value of a patent portfolio may be higher than the sum of the value of individual patents, since patent portfolios might,

for instance, allow companies set standards. The value can only be created if companies have enough resources to monitor patent infringements and subsequently sue the infringers (when patents are infringed it is possible to sue the infringer to recover the profits related to in-house commercialization of products or licensing out, which were lost owing to the infringement).

To use patents effectively (i.e., to create both financial and nonfinancial value), patents can be used in various governance structures: licensing-in, licensing-out, cross-licensing, and patent pools.

Licensing is typically seen as an agreement between the licensor and licensee, where the former permits the latter to use the property rights in commercial applications. This agreement is usually officially signed and it implies that the licensor remains the proprietor of the rights, whereas the licensee pays fees (royalties) to use the patent.

In terms of exclusivity, licenses can be granted either on an exclusive basis (to a single licensee) or on a nonexclusive basis (to many licensees). Exclusive licenses can be restricted by time, geographical territory or market, or by technological field of usage. An exclusive license provides the licensee with a potentially large market power. This type of license is sometimes demanded by companies who plan to invest a substantial amount of money in product development and/or manufacturing. However, there is usually less risk involved for licensor under nonexclusive agreements because having more than one licensee increases the chances of a final product being successful.

In addition to the exclusivity, licensing agreements have other types of clauses (OECD, 2004). Examples of such clauses are as follows: "requirements to work inventions" clauses to make sure that the licensee really tries to commercialize a technology within a reasonable period of time; "requirements to work invention" to permit the patent owner to stop the contract in cases where technology commercialization never takes place; and "reach-through" clauses to secure a licensor's future revenues from licensed technologies, by including contract royalties on future products developed by using a licensed technology; and "property rights on improvements" as both the licensor and licensee are to obtain the rights on these improvements.

Both licensing-in and licensing-out agreements help companies to share risk or liability and also to avoid or to settle litigations and accelerate the development of their technologies. In addition, licensing-in agreements provide firms with the opportunity to develop technology beyond their expertise. Furthermore, in terms of R&D, licensing in helps save on research costs, while licensing out enables to recover all or part of the money spent on R&D. In addition, licensing-out agreements help generate profit. In the case of licensing-out agreements, companies can also achieve a wider diffusion of their technology, which adds value to existing products and also permits companies to test new applications for both new and proven technologies. Other advantages of these agreements is that they enhance customer loyalty, they match promising technologies with the resources needed to bring them to the market, they help to get around trade barriers, they help set standards and, finally, they build relations leading to further collaboration. In addition, licensing-out agreements give an opportunity for companies to increase the scope of usage of their patents. They also give the licensor a chance to profit from the technology entering a new or a foreign market without investing additional resources.

One of the disadvantages of licensing-in is the dependence on the quality of the patents involved in the agreement. For example, the value of the licensed-in patent can be reduced if this patent is less useful than expected or if a third party is infringing the licensor's patent. Licensing-in often implies sharing the patent with other company/ies and, therefore, can force companies to invest more in product development or in marketing. The licensed technology might also become obsolete rapidly and, therefore, the revenue stream will stop, which, in turn, will make the licensed patent a liability rather than an asset. Moreover, the license might be stopped or not be extended by the licensor, which, of course, results in companies losing their customers, revenue, and even, sometimes, reputation. In the case of licensing-out the licensor very often depends on how well and how quickly the licensee develops the licensed patent into a commercial product and the subsequent

success of this commercial product. Another risk is that the revenue generated from licensing patents out can be sometimes lower than the costs of monitoring the licensee's activities. Very often the patent has to be adapted for the licensee, which might result in higher costs for the licensor and, as a consequence, in lower revenue. Licensing-out agreements also might diminish the licensor's control over standards and specifications and, furthermore, its reputation might be damaged or even destroyed by the actions of the licensee.

Cross-licensing agreements usually exist between two parties, and in some rare cases more parties are involved. Cross-licensing agreements imply that companies exchange selected patents on a royalty-free basis, as, in theory, the patents brought by each company in the agreement have the same value. Generally, companies cross-license all the patents within a field of use for a fixed period (this includes existing patents and those that are still pending). The period of cross-license and residual rights reflect the technology life cycle.

There are two main motivations for cross-licensing: they provide companies with freedom to invent without the risk of patent suits and cut down the transaction costs of identifying and monitoring infringements. This type of agreement can be quite extensive and is based on a contract between the parties not to sue each other for infringing of patents they put in the agreement. The agreement may also include pending patents or any other related patents granted in the future. Basically, a cross-licensing agreement can help solve the problems attributed to a patent thicket, which is, according to Shapiro (2001), an overlapping set of patents requiring multiple licensees for those who plan to commercialize a certain technology. Cross-licensing agreements, in this case, reduce the uncertainty associated with the investments in R&D.

The last form of patent governance is patent pools. Patent pools are defined as follows: "A patent pool involves a single entity [. . .] that licenses the patents of two or more companies to third parties as a package" (Shapiro, 2001). Patent pools are characterized by multiparty ownership. They may include either only current patents or also any future changes in these patents. A typical pool makes all the pooled patents available to each member of the pool. Pools also usually offer standard licensing terms to licensees who are not members of the pool. In addition, a typical patent pool allocates a portion of the licensing fees to each member according to a preset formula or procedure.

The basic economic rationale of patent pools is that they significantly reduce the transaction costs of exchanging rights, when compared to a series of bilateral licensing deals. When patent pools are not used as a cover for a cartel, they increase significantly the efficiency of the patent system, as many industries have discovered over time. Department of Justice and Federal Trade Commission (1995) in the United States assert that patent pools need not be open to all who would like to join unless (i) excluded firms cannot effectively compete in the relevant market for the goods incorporating the licensed technologies and (ii) the pool participants collectively possess market power in the relevant market.

The advantages and disadvantages of patent pools overlap with those of cross-licensing agreements. Participation in both types of agreements results in considerable financial savings, through both reduction of royalty stacking and reduction of transaction costs. Some other advantages are related to the creation of relationships between partners. Both governance structures, especially patent pools, are very useful in setting standards. Other advantages are attributed to the sharing and coordinated launch of technology and also clearance of other companies' blocking position. The disadvantage of both cross-licensing agreements and patent pools is believed to be such that agreements might discourage investment in R&D and that members might defect and breach the contract.

Cross-licensing agreements are often more favorable than patent pools because the agreements usually do not involve royalties. In the former, when royalties are involved, they are very often easy to negotiate, which is rarely the case with patent pools as they are multiparty agreements and require a lot of negotiations. As for patent pools, their ultimate advantage lies in the fact that they include a lot of partners at the same time and, in this sense, offer access to a greater amount of skills and resources.

A specific problem related to patent pools is that the great number of patent owners may make it difficult to administrate the pool, to coordinate its development, to negotiate, and to allocate royalties. Patent pool administration costs should not offset transaction savings. A further difficulty is that not all the patents included in patent pools are essential. The fact that only essential patents should be included in the portfolio is the most important feature of a patent pool because it helps ensure that the portfolio patents are not competing with each other. Essentially, all of the patents included in a pool must be the ones that can be used to block the implementation of the standard embodied in the patent pool. It should then lower the costs of the manufacturers who need access to these patents, because prospective licensees will be able to negotiate just once, rather than many times, for all required patents. On the other hand, patent pools may increase the costs of licensees if they include patents that would not otherwise be needed by the licensees.

The Bayh-Dole Act in the United States in 1980s, which made it possible for the United States universities to commercialize their patents, resulted in universities joining the patent race. Similarly, universities in other countries followed the same pattern. Commercialization of university patents is controversial, since depending on the legal system, patents can either belong to universities or to the researchers. Also, university patents usually are results of publicly funded research so there is a debate as to whether it is acceptable to privately appropriate the returns of public investment.

Conclusion

The absence of common IPR law means that companies need to protect their IPR in different regions and even in different countries, which is a luxury for small companies. IPRs, therefore, are only valuable when companies have means to monitor infringement and take the required legal actions. Attempts are now being made by the regulatory authorities to ensure that IPR serve their original purpose of encouraging innovation. For instance, the Adelphi Charter was created in 2005 to address the ever-expanding nature of intellectual property (IP), which puts the future development of creativity and innovation under question.

By establishing a culture of identifying, cultivating, and strategically using their IP assets, companies can increase their revenue, have an edge over their competitors, and position themselves well in the market. A company, however, should focus their resources on inventing, rather than litigating over their innovation, as IP is a necessary but not sufficient ingredient for a company's success.

Bibliography

Department of Justice and Federal Trade Commission (1995) Antitrust Guidelines for the Licensing of Intellectual Property.

Gans, J. and Stern, S. (2003) The product market and the market for "ideas": commercialisation strategies for technology entrepreneurs. *Research Policy*, **32**, 333–350.

Grindley, P. and Teece, D.J. (1997) Managing intellectual capital: licensing and cross-licensing in semiconductors and electronics. *California Management Review*, **39** (2), 8–41.

Kalamas, J., Pinkus, G. and Sachs, K. (2002) The new math for drug licensing. *The McKinsey Quarterly*, **4**, 9–12.

Mazzoleni, R. and Nelson, R. (1998) The benefits and costs of strong patent protection: a contribution to the current debate. *Research Policy*, **27**, 273–284.

Merges, R. and Nelson, R. (1990) On the complex economics of patent scope. *Columbia Law Review*, **90**, 839–916.

OECD (2004) Patent and Licensing at Public Research Organisations. OECD, Paris.

Raustiala, K. and Sprigman, C. (2006) The piracy paradox: innovation and intellectual property in fashion design. *Virginia Law Review*, **92** (8), 1687–1777.

Rayna, T. and Striukova, L. (2008a) Digital rights management: White knight or Trojan horse? *Communication and Strategies*, **69** (1), 109–125.

Rayna, T. and Striukova, L. (2008b) Privacy or piracy, why have to choose? two solutions to the issue of digital rights management and protection of personal information. *International Journal of Intellectual Property Management*, **2** (3), 240–252.

Reitzig, M. (2003) What determines patent value? Insights from the semiconductor industry. *Research Policy*, **32**, 13–26.

Rivette, K. and Kline, D. (2000) *Rembrandts in the Attic: Unlocking the Hidden Value of Patents*, Harvard Business School Press, Boston.

Scotchmer, S. (1991) Standing on the shoulders of giants: cumulative research and the patent law. *Journal of Economic Perspectives*, **5**, 29–41.

Shapiro, C. (2001) Navigating the patent thicket: cross licenses, patent pools and standard-setting, in *Innovation Policy and the Economy*, vol. **1** (A. Jaffe, J. Lerner, and S. Stern), MIT Press for the NBER, Cambridge, pp. 119–150.

Smith, G.V. (1997) *Trademark Valuation*, John Wiley & Sons.

Voice of the Customer

Steven P. Gaskin, Abbie Griffin, John R. Hauser, Gerald M. Katz, and Robert L. Klein

Definition

The *Voice of the Customer* (VOC) is a term used in business to describe the process of capturing customers' requirements. The VOC is a product-development technique that produces a detailed set of customer wants and needs, which are organized into a hierarchical structure, and then prioritized in terms of relative importance and satisfaction with current alternatives.

The VOC process has important outputs and benefits for product developers. VOC provides

- a detailed understanding of the customer's requirements
- a common language for the team going forward
- key inputs for the setting of appropriate design specifications for the new product or service
- a highly useful springboard for product innovation.

There are four aspects of the VOC – customer needs, a hierarchical structure, priorities, and customer perceptions of performance.

Description and Commentary

VOC studies typically consist of both qualitative and quantitative market-research steps. They are generally conducted at the start (or "fuzzy front end") of any new product, process, or service design initiative to better understand the customer's wants and needs (*see* FRONT END OF INNOVATION). The VOC can also be a key input for new-product definition, QUALITY FUNCTION DEPLOYMENT (QFD), or the setting of detailed design specifications (*see* PRODUCT SPECIFICATIONS). It is critical that the product development core team own and be highly involved in this process. They must be the ones who take the lead in defining the topic, designing the sample (i.e., the types of customers to include), generating the questions for the discussion guide, either conducting or observing and analyzing the interviews, and extracting and processing the needs statements. Only by being highly involved can the team fully internalize the VOC and make effective product-design decisions (*see* PRODUCT DESIGN).

As noted in the definition, there are four aspects of the VOC – customer needs, a hierarchical structure of the needs, priorities, and customer perceptions of performance.[1]

Customer needs. A customer need is a description, in the customer's own words, of the benefit to be fulfilled by the product or service. For example, when describing diagonal lines on a computer monitor, a customer might want them "to look like straight lines with no stair-step effect." Note that *the customer need is not a solution*, such as a particular type of monitor (XGA, Megapixel, flat screen, flat panel, etc.), nor a physical measurement (number of noticeable breaks in the line), but rather a detailed description of how the customer wants images to appear on the monitor (Griffin and Hauser, 1993).[2]

The distinction between physical measurements and customer needs has proven to be one of the keys to the success of marketing tactics. As illustrated in Figure 1, the "lens" model suggests that customers see the world through the lens of their perceptions (their needs) (Brunswick, 1952). The lens model says that customers choose (buy a product or service) if they prefer that product over others and it is available to them in the marketplace. However, preferences are based on how customers perceive the world. This perception may or may not be totally accurate. It is based, of course, on the product's features, but it is also based on the image created by advertising, packaging, word of mouth, social context, and so on. Marketing is an integrated activity that attempts to design the product (physical features) and the marketing to influence customer perceptions. Within the context of the lens model, the VOC identifies the dimensions of customer value (customer needs) and how customers form preferences with respect to those needs (importance of those customer needs). The VOC might also identify how advertising, and so on, affect perceptions, availability, and perceived price.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

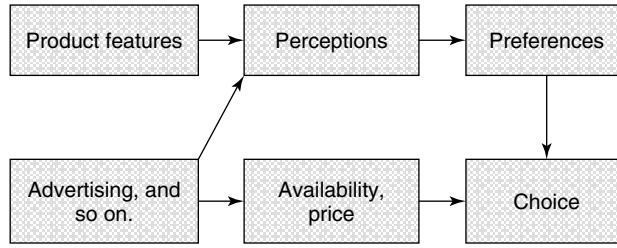


Figure 1 The lens model of customer choice.

Knowing customer needs is critical to both product development and marketing. For example, if a product-development team focuses too early on solutions, they might miss creative opportunities. A computer-monitor team might be tempted to focus on the size of the monitor or the shape. However, readability might also depend on the ambient room light and reflections, the colors that the software designer chooses, the ratio of the height of small letters to that of capital letters, and even the style of the typeface (serif or sans serif, proportional or fixed, etc.). All of these design attributes interact with the size and shape of a monitor to affect the customer need of "easy-to-read text." Some may be less costly and more effective, some may be synergistic with changing the monitor's size and shape, but all should be considered before a final design is chosen for the monitor.

Discussions with customers usually identify 75–150 phrases that might be considered an articulation of customer needs. Such phrases might include basic needs (what a customer assumes a monitor will do), articulated needs (what a customer will tell you that he, she, or they want a monitor to do), and exciting needs (those needs which, if they are fulfilled, would delight and surprise the customer) (*see also* KANO MODEL OF CUSTOMER SATISFACTION). It is extremely important that these customer needs be stated *in the customers' own words*, and not in industry or company jargon, in order to not lose the meaning.

Identifying customer needs is primarily a qualitative research task. In a typical study, between 10 and 30 customers are interviewed for approximately one hour in a one-on-one setting. For example, a customer might be asked to picture him- or herself viewing work on a computer. As the customer describes his or her experience, the interviewer keeps probing, searching for better and more complete descriptions of how he or she views data, images, video, or anything else, how he or she works with those images, working conditions, ambient lighting, and so on. The goal is to experience the experience of the customer. Sometimes, the interviews take place at the site where the customer uses the product – for example, VOC interviews have been conducted on oil-drilling platforms for manufacturers of oil-drilling equipment. This method of data collection is sometimes referred to as *customer visits* (McQuarrie, 2008), *contextual inquiry*, or *ethnography*.

The interviews are called "*experiential*," because they focus on the customers' experiences. In the interview, the customer might be asked to voice needs relative to a number of real experiences. The interview ends when the interviewer feels that no new needs can be elicited from that customer.

While it is tempting to simply ask customers, "What are your needs?" customers often have difficulty articulating them. It is much better to infer customer needs from experiential interviews or observation.

Hierarchical structure. The average marketing manager cannot work directly with the 75–150 detailed customer needs found in the first step of the VOC process. A simpler structure is needed that focuses both strategy and tactics. The "Voice of the Customer" structures customer needs into a hierarchy of primary, secondary, and tertiary needs. *Primary needs*, also known as *strategic needs*, are the 2–10 top-level needs that are used by the team to set the strategic direction for marketing. Each primary need is elaborated

into 3–10 secondary needs. *Secondary needs* indicate more specifically what the marketing manager must do to satisfy the corresponding primary (strategic) need. (Secondary needs are also known as *tactical needs*.) *Tertiary needs*, also known as *operational* or *detailed needs*, provide greater detail so that engineering, R&D, and, perhaps, the advertising agency, can develop a detailed set of product characteristics or advertising copy that satisfies the primary and secondary needs.

For example, a VOC analysis for movie theaters identified the following 17 secondary (tactical needs) structured into 6 primary (strategic needs):

- Theater selection
  - Offers a good selection of movies and show times
  - Easy to get information about show times
  - I can always get into the movie I want to see
  - A variety of easy and economical ways to buy tickets
- Getting to the theater
  - The theater is conveniently located
  - There is safe, convenient parking
- Food/refreshments
  - Good food is available at a fair price
  - Concessions are well run
- The theater building
  - Quick and easy access to everything I need
  - Handles crowds well
  - Friendly and available customer service
  - Clean, well-equipped restrooms
  - A comfortable feeling inside the theater
- Inside the theater auditorium
  - Clean, comfortable seating
  - Auditoriums are clean
- The movie experience
  - A great view and sound so I'm right in the action
  - No disturbances during the show.

There are a number of ways to group the 75–150 tertiary (operational) needs into a more aggregate set of tactical needs, and then further into an even smaller set of strategic needs. The easiest way is to have the product-development team do so as a group. However, while convenient, this approach has the important limitation that the results tend to reflect the company's organizational chart, that is, how the product is developed and produced, rather than the way customers think. It is far better to have customers group the needs. One way this can be done is through the use of one or a few focus groups. A moderator guides the process and makes sure the groupings make sense and are sufficiently disaggregate (i.e., that there are sufficiently many tactical needs so that they do not cover multiple topics). A more statistically representative method is to survey a random sample of current and potential customers and have them sort the needs individually into piles based on similarity. This results in a co-occurrence matrix, which can be analyzed using a hierarchical clustering routine. The output is a dendrogram, which shows how the needs should be grouped for any total number of needs, from the total number of detailed needs down to two or more strategic needs. The final number of primary and secondary needs are then determined judgmentally on the basis of the output of the cluster analysis.

Priorities. Some needs have higher priorities for customers than others. The marketing manager uses these priorities to make decisions that balance the cost of fulfilling a customer need with the desirability (to the customer) of fulfilling that need. For a movie theater company, for example, the strategic decision on whether to provide or communicate improved movie-viewing experience depends upon the cost and feasibility of improving the experience and the priority to the customer of an improved viewing experience relative to the customer's other needs. In the VOC, these priorities apply to perceived customer needs rather than product features or engineering solutions. As an example, a quantitative survey of people who go to movie theaters yielded the following importance weights for the 17 secondary needs (Figure 2).

A number of new techniques have been developed recently to prioritize customer needs. One, called *teaching agents to choose*, is an incentive compatible direct-elicitation technique.[3] During the survey, respondents

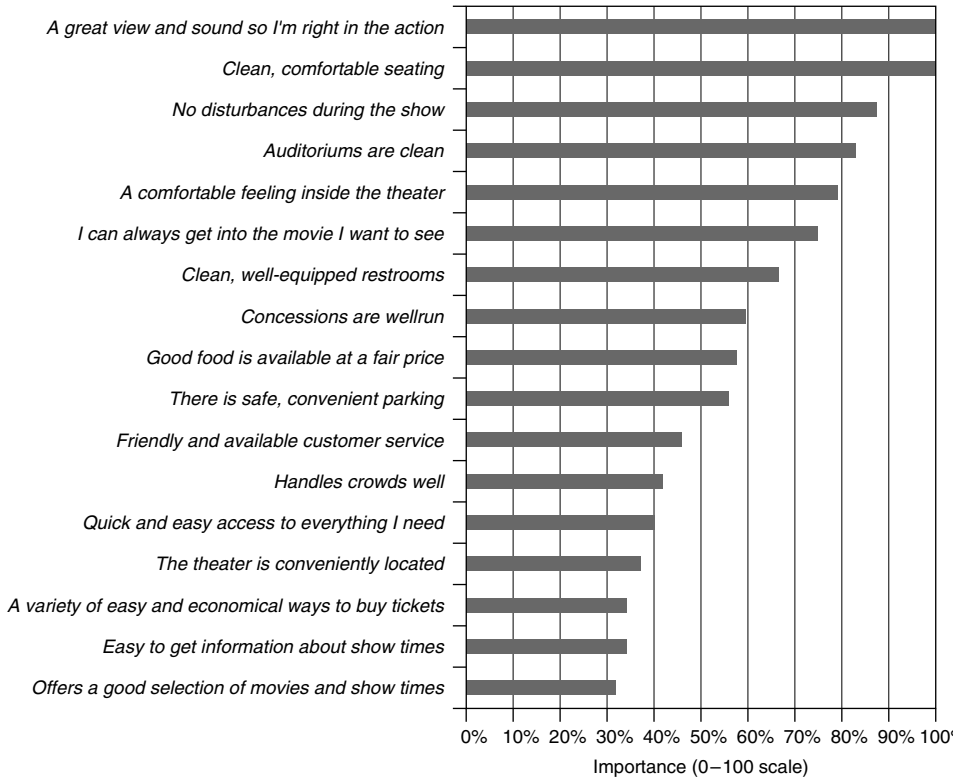


Figure 2 Importance ratings for movie theater needs.

give instructions to a hypothetical "agent" about what rules to use when deciding to choose which product to buy. Priorities of the needs can be estimated by tallying the frequency of mention of the needs referenced in these rules. Another technique, MaxDiff, involves an exercise where respondents pick the most and least important of a series of subsets of needs. The resulting scores are not subject to scale usage bias, which can lead to problems, particularly in international studies.

Customer perceptions of performance. Customer perceptions are also derived from quantitative market research about how customers perceive the performance of products that compete in the market being studied. If no product exists as yet, the perceptions indicate how customers now fulfill those needs. (For example, existing patterns of medical care serve as generic competition for health-maintenance organizations. Automobile and bus transportation serve as generic competition for Southwest Airlines.) Knowledge of which products fulfill which needs best, how well those needs are fulfilled, and whether there are any gaps between the best product and "our" existing product provide further input into marketing decisions.

Customer perceptions are often displayed via a "snake plot," called so because each product's performance "snakes" across the page. These data are often obtained via a questionnaire in which each respondent rates each product (that they consider) on each of the secondary customer needs.

For example, when people who go to movie theaters were asked to rate how well the theater they attend most often performed on

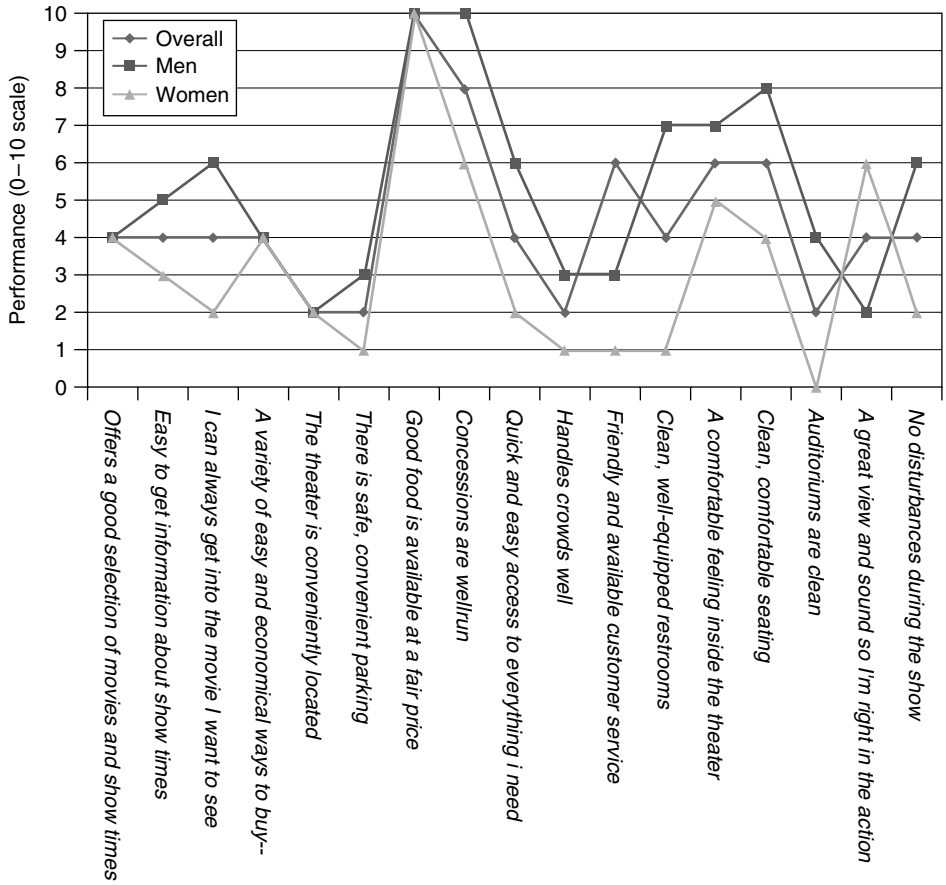


Figure 3 Performance ratings for movie theaters.

the secondary needs, we obtained the plot in Figure 3.

Examples of VOC's Role in Product Development

During a VOC for the renal division of Baxter Healthcare, users expressed a need to quickly understand the nature and seriousness of the various alarms that typically sounded throughout the day. The product's designers developed a traffic-light device with colored lights that indicated the gravity of the problem at a glance.

A few years ago, Intel, the world's largest manufacturer of microprocessors, embarked on a very public strategy to develop specialized microprocessors for various applications. To better understand customer needs surrounding these applications, the company began to develop its own VOC. Using a combination of ethnography and sit-down interviews, followed by group affinitization processes and large-scale quantitative prioritization surveys, the company has now conducted more than 20 such studies on a host of topics, ranging from extremely simple consumer applications, to highly advanced applications with IT Directors at very large companies.

One of the most compelling products to emerge from this strategy is the v-Pro™ microprocessor for business desktop computers. The v-Pro™ technology addresses a number of critical problems for IT Directors who must manage a large "fleet" of company computers, and

ensure their security from external corruption and tampering. The v-Pro™ offers several new features – which have proven to be extremely attractive to IT Directors – that emerged from the identification of key unmet needs in the VOC process. Launched in mid-2006, the v-Pro™ became the fastest product in Intel's history to exceed $1 billion in revenue.

PPG Industries, a leading manufacturer of industrial coatings and other commercial materials, trained product development teams across most of its major divisions in the use of VOC. These teams then conducted a number of highly successful VOC studies throughout the world, on topics as diverse as applications for polyurethane coatings, auto paint used in body shops, and eco-friendly uses for fiberglass. Products that emerged from these applications include a chemical-agent-resistant coating used in military applications, a new material used for golf-ball covers, and an advanced fiberglass product used in wind-power generation. This latter product allowed PPG to acquire a major share of this rapidly expanding industry.

As may be seen in these examples, gathering the VOC is an extremely important part of the "fuzzy front end" of the new-product development process. It forms a solid basis for design and marketing decisions from concept development through product launch.

Endnotes

[1] The Voice of the Customer has its origins in the QFD process, where it is used to develop the customer needs that are linked to performance measures. This is why the definition used here is narrower than the generic use of VOC, which can refer to customer feedback in any form. See Griffin and Hauser (1993).

[2] Much of the material used here is drawn from an MIT Sloan Courseware document by John R. Hauser, "*Note on the Voice of the Customer*," MIT, Cambridge, MA 2008. MIT Sloan grants the nonexclusive right to use this material. MIT Sloan retains a nonexclusive right to this material.

[3] "Incentive compatible" means the incentives for survey respondents are set up in a way that they benefit most from responding in a desired way (in this case, telling the truth).

Bibliography

Brunswick, E. (1952) *The Conceptual Framework of Psychology*, University of Chicago Press, Chicago.

Hauser, J.R. and Clausing, D.P. (1988) The house of quality. *Harvard Business Review*, **66** (3), 63–73.

Griffin, A. and Hauser, J.R. (1993) The voice of the customer. *Marketing Science*, **12** (3), 1–27.

Katz, G. (2001) The "One Right Way" to gather the voice of the customer. *PDMA Visions*, **25**.

Katz, G. (2004) The voice of the customer, in *The PDMA Toolbook 2 for New Product Development*, Chapter 7 (eds P. Belliveau, A. Griffin, and S. Somermeyer), John Wiley & Sons, Inc., Hoboken.

McQuarrie, E.F. (1998) *Customer Visits*, Sage Publications, Thousand Oaks.

McQuarrie, E.F. (2008) *Customer Visits*, M.E. Sharpe, New York.

success factors for new-product development

Robert G. Cooper and Elko J. Kleinschmidt

Introduction

Product innovation – the development of new and improved products – is crucial to the survival and prosperity of the modern corporation. New products currently account for up to 33% of company sales (Griffin, 1997; American Productivity & Quality Center, 2003), while product life cycles (*see* PRODUCT LIFE CYCLE; TAKEOFF) are getting shorter: a 400% reduction over the last 50 years, the result of an accelerating pace of product innovation. But many products do not succeed: approximately one in ten product concepts succeeds commercially and only 51% are launched on time (American Productivity & Quality Center, 2003). Further, only one in four development projects is a commercial success (Cooper, 2001), while companies scrap almost one-third of new-product projects incurring losses of $80 billion annually.

The huge amounts at stake coupled with the high odds of failure make product innovation one of the riskiest endeavors of the modern corporation. Thus many managers, researchers, and pundits have sought answers to the age-old question: what makes a new product a winner? And why are some businesses so much more successful at product development than the rest?

We look at the critical success factors that distinguish new-product winners. These factors are based on studies that focus on individual new-product projects (factors 1–7 below); and also on studies that assess the business's overall new-product performance (factors 8–13 below).

What Makes New Products Successful

Factor 1: A unique superior product. A superior and differentiated product is the number one driver of success and profitability, with success rates reported to be three to five times higher than for "me too," reactive products (Cooper, 2001; Cooper and Kleinschmidt, 1996). Well-known studies (NewProd Studies by Cooper *et al.*, see Cooper, 2001) show the dramatic impact of product superiority: For example, such products (the top 20%) when compared to those with the least degree of differentiation (the bottom 20%)

- have an exceptional commercial success rate of 98.0%, versus only 18.4% for undifferentiated ones;
- have a market share of 53.5% of the defined target market, versus only 11.6% for "me too" new products;
- have a rated profitability of 8.4 out of 10 (versus only 2.6 out of 10 for undifferentiated products; here 10 = exceptional profits, far exceeding the company's minimum hurdle);
- meet company sales and profit objectives to a greater degree than do undifferentiated products.

The same studies show, however, that reactive products, undifferentiated products, and technically driven products that lack customer benefits are the rule rather than the exception; and the majority fail.[1]

What do these superior products with unique customer or user benefits have in common? These winning products

- feature good value for money for the customer, reduce the customer's total costs (high value in use), and boast excellent price/performance characteristics;
- provide excellent product quality relative to competitors' products, in whatever manner the user measures quality (*see* COMPETITIVE ADVANTAGE);
- are superior to competing products in terms of meeting users' needs, offer unique features not available on competitive products, or solve a problem the customer has with a competitive product;
- offer product benefits or attributes easily perceived as useful by the customer, and benefits that are highly visible (*see* VALUE PROPOSITION).

Note that the product must be superior in the eyes of the customer, not just the eyes of the engineer, scientist, or designer: often developers heavily overestimate the customer benefits and desirability of their products (Cooper, Edgett, and Kleinschmidt, 2003).

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Factor 2: A strong market orientation – market driven, customer focused. A thorough understanding of customers' needs and wants, the competitive situation, and the nature of the market is an essential component of new-product success (*see* VOICE OF THE CUSTOMER). This finding is supported in virtually every study of product success factors (Cooper, 2001). Recurring success themes include

- need recognition
- understanding user needs
- market-need satisfaction
- constant customer contact
- strong market knowledge and market research
- quality of execution of marketing activities
- more spending on the up-front marketing activities.

Even in the case of technology-driven new products (where the idea comes from a technical or laboratory source), the likelihood of success is greatly enhanced if customer and marketplace inputs are built into the project soon after its inception.

Conversely, a failure to adopt a strong market orientation in product innovation, an unwillingness to undertake the needed market assessments and to build in the Voice of the Customer, and leaving the customer out of product development spell disaster. Poor market research, inadequate market analysis, weak or no market studies, test markets, and market launch, and inadequate resources devoted to marketing activities are common weaknesses found in almost every study of why new products fail (Cooper, 2001).

A strong market orientation must be built into every stage of the new-product process to achieve success (Cooper, 2001):

- *Idea generation*: focusing on the customer, and using the customer as a source of ideas (*see* IDEA MANAGEMENT).
- *Product design*: employing market research as an input to product design, not just an after-the-fact check (*see* PRODUCT DESIGN).
- *During development*: maintaining constant customer contact and feedback (e.g., continuous customer testing of facets of the product) (*see* PRODUCT TESTING).
- *After development*: undertaking customer trials, preference tests, and test markets to verify market acceptance and launch plan (*see* PRETEST MARKET MODELS).
- *Launch*: employing a well-designed, carefully targeted, properly resourced launch, guided by a well-conceived marketing plan based on solid market information (*see* LAUNCH STRATEGIES).

Factor 3: The world product – a global orientation in product development. The world is the business arena today, and thus company growth and profitability are dependent on strategies of globalization and new-product development (*see* GLOBAL PRODUCT DEVELOPMENT). In global markets, new-product development plays a primary role in achieving a sustainable competitive advantage (Kleinschmidt, de Brentani, and Salome, 2007).

A recent study found that multinational firms that take a global approach to new-product development outperform those that concentrate their research spending in their home market (The Economist, 2008, November 22–28, p. 74). Further, an investigation by de Brentani and Kleinschmidt (2004) and Kleinschmidt, de Brentani, and Salome (2007) also showed that firms that are strongly globally minded in their new-product development perform significantly better. International products targeted at world and nearest-neighbor export markets are the top performers (Cooper, 2001). By contrast, products designed for only the domestic market, and sold to domestic and nearest-neighbor export markets fare more poorly. The magnitude of the differences between these international and exported new products versus domestic products is striking: differences of two- or three-to-one on various performance gauges.

Increasing globalization of markets demands global products. To define the new-product market as "domestic" and perhaps a few other "nearby convenient countries" severely limits market opportunities. For maximum success in product innovation, the objective must be to design for the world and market to the world. Sadly, this international dimension is an often

overlooked facet, or one which if included, is handled late in the development process or as a side issue (*see also* LOCALE OF INNOVATION).

An international orientation means defining the market as an international one, and designing products to meet international, not just domestic, requirements. The result is either a global product (one version for the entire world) or a glocal product (one development effort, one product concept or platform, but perhaps several product variants to satisfy different international markets). An international orientation also means undertaking market research, concept testing, and product testing in multiple countries, rather than just the home country; and it means relying on an international or global project team (only one new-product project team in five is reported to be a global development team) (Cooper, 2001).

Factor 4: Predevelopment work – the up-front homework. The steps that precede the design and development of the product make the difference between winning and losing (Cooper, 2001). New-product projects that feature a high quality of execution of activities that precede the development phase – the fuzzy front end – are more successful:

- a success rate of 75.0% (versus only 31.3% for projects where the predevelopment activities are found lacking);
- a higher rated profitability (7.2 out of 10 versus only 3.7 for projects where predevelopment activities are poorly undertaken);
- a market share of 45.7% (versus 20.8%) (Cooper, 2001).

Successful businesses spend about twice as much time and money on these vital up-front or predevelopment activities, such as initial screening, preliminary market and technical assessments, detailed market studies or marketing research (success factor 2), and business and financial analysis just before the decision to "Go to Development" (Cooper, 2001).

Most businesses confess to serious weaknesses in the up-front or predevelopment steps of their new-product projects. Small amounts of time and money are devoted to these critical steps: only about 7% of the project's expenditure and 16% of the effort. Far from adding extra time to the project, research reveals that homework pays for itself in reduced development times, the result of sharper and more stable product definition, and fewer surprises (and time wasters) later in the project (Cooper, 2001).

Factor 5: Sharp and early product definition. Successful products have much sharper definition prior to development. Products that have these sharp definitions are 3.3 times as likely to be successful and have higher market shares (by 38 share points on average); they are much more profitable than poorly defined products, and they meet company sales and, more so, profit objectives (Cooper, 2001).

This definition includes

- specification of the target market: exactly who the intended users are;
- description of the product concept and the benefits to be delivered;
- delineation of the positioning strategy (including the target price) (*see* PRODUCT POSITIONING);
- and a list of the product's features, attributes, requirements, and specifications (*see* PRODUCT SPECIFICATIONS).

Unless these four items are clearly defined, written down, and agreed to by all parties prior to entering the development stage, the odds of failure increase by a factor of 3, for the following reasons (Cooper, 2001):

- Building a definition step into the new-product process forces more attention to the up-front or predevelopment activities.
- The definition serves as a communication tool and guide. All-party agreement or "buy in" means that each functional area involved in the project has a clear and consistent definition of what the product and project are, and is committed to it.
- This definition also provides a clear set of objectives for the development phase of the project, and the development team members. With clear product objectives, development typically proceeds more efficiently and quickly: there are no moving goalposts and no fuzzy targets.

Factor 6: The way project teams are organized. Product innovation is very much a team effort. On doing a post mortem on any bungled new-product project, invariably one will find each functional area doing its own piece of the project, with very little communication between players and functions – a fiefdom mentality; and no real commitment of project team members to the project. Many studies concur that how the project team is organized and functions strongly influences project outcomes (*see also* ORGANIZING FOR INNOVATION).

Product development must be run as a multidisciplinary, cross-functional effort (*see also* CROSS-FUNCTIONAL TEAM). Good organizational design means projects that are

- organized as a cross-functional team with members from R&D, engineering, marketing and sales, operations, and so on (as opposed to each function doing its own part independently);
- where the team is dedicated and focused (i.e., devotes a large percentage of their time to this project, as opposed to spread over many projects or other work);
- where the team members are in constant contact with each other, via frequent but short meetings, interactions, project updates, and even colocation;
- where the team is accountable for the entire project from beginning to end (as opposed to accountability for only one stage of the project);
- where there is a strong project leader or champion who leads and drives the project (*see also* LEADERSHIP ROLES IN PRODUCT DEVELOPMENT).

While the ingredients of good organizational design should be familiar, surprisingly, many firms have yet to get the message (American Productivity & Quality Center, 2003; Cooper and Kleinschmidt, 1996).

Factor 7: Leveraging core competencies. Leverage and synergy is the common thread that binds the new business to the old. When translated into product innovation, the ability to leverage existing and in-house strengths, competencies, resources, and capabilities increases the odds of success of the new-product project (*see* CORE COMPETENCIES). By contrast, "step-out" projects take the firm into territory that lies beyond the experience and resource base of the company, and increase the odds of failure (Cooper, 2001).

The reasons for the impact of leverage are clear:

1. Resources are available and at marginal cost.
2. Operating within one's field of expertise – either familiar markets or familiar technologies – provides considerable "domain knowledge," which is available to the project team.
3. The more often one does something, the better one becomes at doing it – the experience factor. If new-product projects are closely related to (leveraged from) current businesses, the chances are, there has been considerable experience with such projects in the past, hence lower costs of execution and fewer misfires.

Two types of leverage are important to product innovation:

- *Technological leverage*: the project's ability to build on in-house development technology, use inside engineering skills, and use existing manufacturing or operations resources and competencies.
- *Marketing leverage*: project/company fit in terms of customer base, sales force, distribution channels, customer-service resources, advertising and promotion, and market-intelligence skills, knowledge, and resources.

These two dimensions of leverage – technological and marketing, and their ingredients – become obvious checklist items in a scoring or rating model to help prioritize new-product projects. If leverage is low, yet the project is attractive for other reasons, then steps must be taken to bolster the internal resources and competencies. Low leverage scores signal the need for outside resources – partnering or outsourcing. But neither solution is a panacea: there are risks and costs to both routes to securing the needed resources and competencies (Campbell and Cooper, 1999).

Innovation versus incremental product development. The term *new products* encompasses a wide range of development projects, from product improvements to true innovations (*see* INNOVATION TYPOLOGIES). Note that most "new products" are not truly new to the world products or breakthroughs (*see* RADICAL INNOVATION); rather they are incremental innovations such as cost reductions, improvements, product-line extensions, and new items in existing product lines. Such incremental products, often called *sustaining innovation*, are necessary to keep the firm's product portfolio current and competitive, and are based on exploiting existing company knowledge, markets, and customers; not surprisingly, they represent the great majority (approximately 90%) of all new projects (Cooper, 2001; Cooper, Edgett, and Kleinschmidt, 2003).

Seeking success by applying some of these seven success factors above could lead to a "success trap." For example, accumulated decision-making experiences will favor incremental projects over true innovations (March, 1991; Levinthal and March, 1993; Teece, Pisano, and Shuen, 1997). Additionally, rigorously applying factors 2, 5, and 7 – a strong market orientation, seeking sharp, early product definition, and leveraging core competencies – could lead to picking closer-to-home and less risky development projects. Thus, employing factors 1–7 will certainly increase new-product success rates, but some factors, when make it difficult to achieve real breakthroughs that are aimed at new markets, may require new technical and market knowledge; these factors are typically difficult to define in the early stages. The result could be high success rates, but lower overall new-product performance for the firm – a "win the battle, but lose the war" outcome.

What Distinguishes the Top-Performing Businesses

A number of success factors have been uncovered by research into what makes some businesses more successful at product innovation. Factors 8–13 below are some of the more important factors that separate the top-performing businesses from the rest in product development.

Factor 8: The right organizational culture. There is ample research-based evidence that a strong innovation culture has a positive impact on performance. Characteristics of this culture include (de Brentani and Kleinschmidt, 2004; Kleinschmidt, de Brentani, and Salome, 2007)

- managers and employees that believe in the importance of new products for the company's continued success, domestically and internationally;
- a mind-set that motivates employees to holistically endorse a belief in creating newness;
- company behavior that is comfortable with, even aggressive about, new ideas, change, risk, and failure;
- an atmosphere where entrepreneurship and risk taking are encouraged and rewarded;
- individuals or teams that are not inadvertently punished when new products do not achieve expected results;
- openness and informal communication and involvement, thinking outside the box, and adaptability to change, as preferred modes of operation.

The best companies in the APQC best practices study (2003) emphasize employee recognition for new-product performance (Griffin, 2008). A positive culture also means senior managers refraining from "micro-managing" projects and second-guessing project team members; and resources and time being made available for creative people to work on their own "unofficial projects," for example, via free scouting-time or bootstrapping funds; and permitting skunk works projects outside the official bureaucracy. Idea-submission schemes (where employees are encouraged to submit new product ideas), and open project review meetings (where the entire project team participates), are other facets of a positive climate (Cooper, Edgett, and Kleinschmidt, 2003). Most businesses, however, are quite weak on almost all these ingredients of a positive culture or climate, with typically less than one-third of businesses employing these practices. By contrast, top-performing businesses embrace these practices to a greater extent.

Factor 9: A product innovation and technology strategy for the business. We live in turbulent times. Technology advances at an ever increasing pace; customer and market needs are constantly changing; competition moves at lightning speed; and globalization brings new players and opportunities into the game. More than ever, businesses need a product innovation and technology strategy to help chart the way (Cooper, 2000) to achieve success (de Brentani, Kleinschmidt, and Salomo, 2009).

Having a new-product strategy for the business is strongly linked to positive performance (American Productivity & Quality Center, 2003). Top-performing businesses in product development clearly define new-product goals for the business; for example, what percentage of the business's sales, or profits or growth will come from new products (*see also* OPPORTUNITY IDENTIFICATION). And the role of product innovation in achieving overall business goals is clearly articulated. Most important, an innovation strategy has a longer time horizon and commitment; surprisingly, only 38.1% of businesses having a longer term new-product strategy, but most top performers do (American Productivity & Quality Center, 2003).

Top-performing businesses focus their R&D efforts by delineating strategic arenas: strategic areas on which the business will concentrate its new-product efforts (*see* RESEARCH & DEVELOPMENT). Defined arenas might include industry sectors, specific markets, product types, or technology types.

Strategic buckets is a method applied by top performers to ensure strategic alignment of product innovation with business goals (*see* PORTFOLIO MANAGEMENT). The strategic-buckets method dedicates resources (funds or person days) to different project types, for example, incremental versus innovative projects, or across strategic arenas, thereby fostering the right mix and balance of development projects.

Finally, strategic product road maps are also employed by top performers. A road map is simply a management group's view of how to get where they want to go or how to achieve their desired objective. The road map portrays a series of major development initiatives – product and technology developments – sequenced over time, typically five to seven years into the future.

Factor 10: Resource commitment. Too many projects simply suffer from a lack of time and money commitment. The results are predictable: much higher new-product failure rates (Cooper, 2001; Cooper and Kleinschmidt, 1996) and poorer new-product performance for the business overall (de Brentani and Kleinschmidt, 2004). Some facts, in this context, are listed below:

- A strong market orientation is missing in the typical new-product project; and much of this deficiency is directly linked to a lack of marketing resources available for the project.
- Another serious pitfall is that the front-end homework does not get done. Again much of this deficiency can be directly attributed to a lack of resources: simply not enough money, people, and time to do the work.

The reason is that as the competitive situation has toughened, companies have responded with restructuring and doing more with less. Consequently, resources are limited or cut. This short-term focus takes its toll. Certain vital activities, such as market-oriented actions and predevelopment homework are highly under-resourced, particularly in the case of product failures.

Best-practice companies commit the necessary resources to new products much more than do most firms (American Productivity & Quality Center, 2003; de Brentani and Kleinschmidt, 2004). While product development resources are in short supply across the board – with less than 30% of businesses indicating "sufficient resources" from four key functional areas – top-performing businesses appear to be much better resourced. Equally important, these resources are focused and dedicated, with project team members not working on too many projects or doing other tasks. Indeed, about half the top performers have a ring-fenced product innovation group that does nothing but work on new products (this is a dedicated cross-functional group – technical, marketing, even sales and operations – whose full-time job is to work on new-product projects, and which typically

reports to a senior executive). Finally, resources are available early in the project to undertake the essential up-front homework and early-stage market research outlined above.

Factor 11: Top management support. Top management support is a necessary ingredient for product innovation and is directly linked to the business's performance in product development (Cooper, 2001; de Brentani and Kleinschmidt, 2004; Kleinschmidt, de Brentani, and Salome, 2007). But it must be the right kind of support. The top management's role in product development is as a facilitator – to set the stage – and not as an actor, front and center.

One important role of senior management is to lead the innovation effort in their business. In other words, they articulate a new-product strategy for the business, something that is often missing. Management in top-performing businesses also commits the necessary product development resources, and keeps the commitment (American Productivity & Quality Center, 2003). They support a disciplined process to drive products to market. These three factors – an articulated new-product strategy, adequate resources, and a disciplined new-product process – are the strongest drivers of a business's new-product performance (Cooper, 2001; Cooper and Kleinschmidt, 1996).

The senior management has a major impact on and fosters a positive climate and culture for innovation within the business, supporting innovative activities and projects at every opportunity with actions, not just words. Finally, the senior management empowers project teams and supports committed champions by acting as mentors, facilitators, "godfathers," or executive sponsors to project leaders and teams – acting as "executive champions."

Factor 12: Tough Go/Kill decision points and better focus (portfolio management). Most companies suffer from too many development projects and not enough resources to mount an effective or timely effort on each, that is, there is a lack of focus. This stems from inadequate project evaluation and poor project prioritization. Project evaluations are consistently cited as weakly handled or nonexistent: decisions involve the wrong people from the wrong functions (no functional alignment); no consistent criteria are used to screen or rank projects; or there is simply no will to kill projects at all (Cooper, Edgett, and Kleinschmidt, 2001; Cooper and Kleinschmidt, 1996).

The desire to weed out bad projects coupled with the need to focus limited resources on the best projects means that tough Go or Kill and prioritization decisions must be made. Some companies have built funnels into their idea-to-launch process via a number of decision points in the form of gates. At successive gate reviews, the senior management rigorously scrutinizes projects, makes Go or Kill and prioritization decisions based on visible Go/Kill criteria, and culls out the weaker or less-valuable projects.

Project selection and picking winning new-product initiatives is only part of the task, however. Other goals are selecting the right mix and balance of projects in the development portfolio, seeking strategic alignment in the portfolio, and ensuring that the business's development spending mirrors its strategic priorities. Many businesses have moved to more formal portfolio management systems (*see* PORTFOLIO MANAGEMENT) to help effectively allocate resources and prioritize new-product projects (Cooper, Edgett, and Kleinschmidt, 2001).

A vital issue in portfolio management is achieving the right mix and balance of development projects in the portfolio. Development portfolios in top-performing businesses contain a higher proportion of more innovative development projects, while portfolios in the typical firm have far too many low-value, noninnovative projects. To foster more innovative development projects, top-performing businesses articulate a product innovation strategy that specifies the quest for true innovations (success factor 9) (American Productivity & Quality Center, 2003). They set aside sufficient resources for innovative development work via strategic buckets – PORTFOLIO MANAGEMENT outlines how to achieve the correct balance between riskier, longer-term and more innovative development projects versus shorter-term, incremental innovations. Additionally, the top management supports such innovative projects and is not afraid to invest in riskier developments

(success factor 11). Finally, top-performing businesses employ flexible development or idea-to-launch processes that accommodate more innovative products, or have adopted quite different processes for such high-risk projects (Cooper, 2006) (THE STAGE-GATE IDEA TO LAUNCH SYSTEM deals with the issue of adaptive, scalable, and flexible development processes in more detail).

Factor 13: A multistage, disciplined new-product process. A systematic new-product process – a Stage-Gate® system[2] – is the solution that many companies have turned to in order to overcome the deficiencies that plague their new-product programs (Cooper, 2001). A Stage-Gate® process is simply a road map or "play book" for driving new products from idea to launch, successfully and efficiently. For major projects, the system consists typically of five stages from Discover though to Launch, with critical best-practice activities laid out for the project team to execute within each stage. Each stage is preceded by a gate or Go/Kill decision point, where an investment decision is made to continue with the project, or to cut one's losses and kill a bad project.

Almost every top-performing company has implemented such a stage-and-gate system to drive their new-product projects through to commercialization, according to the APQC benchmarking study (2003). The payoffs of such processes have been frequently reported: improved teamwork, less recycling and rework, improved success rates, earlier detection of failures, a better launch, and even shorter cycle times (by about 30%) (*see* THE STAGE-GATE IDEA TO LAUNCH SYSTEM for the details on the idea-to-launch or Stage-Gate® process).

Conclusion

Product innovation is a vital task for the modern corporation. The company's success at new-product conception, development, and launch may well decide the fate of the entire business. But many companies miss the mark when it comes to the 13 critical success factors outlined in this article, and the results are predicable – a mediocre new-product effort. These success factors are fact based and proven, and when implemented proficiently, lead to higher success rates at the new-product project level and a higher overall new-product performance for the business.

Endnotes

[1] This material draws on several sources, see Cooper (2001) and Cooper (2003)

[2] The term "*Stage-Gate*" was coined by the R.G. Cooper and first appeared in print in 1988. Stage-Gate is a legal tradename of the Product Development Institute Inc. For more on Stage-Gate® methods, see Cooper (2001); also the Stage-Gate web-page: www.stage-gate.com

Bibliography

American Productivity & Quality Center (2003) Improving New Product Development Performance and Practices, APQC, Houston, www.apqc.org/pubs/NPD2003.

de Brentani, U. and Kleinschmidt, E.J. (2004) Corporate culture and commitment: impact on performance of international new product development programs. *Journal of Product Innovation Management*, **21**, 3009–3333.

de Brentani, U., Kleinschmidt, E.J., and Salomo, S. (2009) Success in global new product development: impact of strategy and the behavioral environment of the firm. *Journal of Product Innovation Management*, forthcoming.

Campbell, A.J. and Cooper, R.G. (1999) Do customer partnerships improve success rates? *Industrial Marketing Management*, **28** (5), 507–519.

Cooper, R.G. (1988) The new product process: a decision guide for managers. *Journal of the Marketing Management*, **3** (3), 238–255.

Cooper, R.G. (2000) "Product Innovation and Technology Strategy" in the "Succeeding in Technological Innovation" series. *Research-Technology Management*, **43** (1), 28–44.

Cooper, R.G. (2001) *Winning at New Products: Accelerating the Process from Idea to Launch*, 3rd edn, Perseus Books, Reading.

Cooper R.G. (2003) Stage-Gate new product development processes: a game plan from idea to launch, in *The Portable MBA in Project Management* (ed. E. Verzuh), John Wiley & Sons, Inc., Hoboken, pp. 309–346.

Cooper, R.G. (2006) Managing technology development projects – different than traditional development projects. *Research-Technology Management*, **49** (6), 23–31.

Cooper, R.G., Edgett, S.J., and Kleinschmidt, E.J., (2001) *Portfolio Management for New Products*, 3rd edn, Perseus Books, Reading.

Cooper, R.G., Edgett, S.J., and Kleinschmidt, E.J. (2003) *Best Practices in Product Development: What Distinguishes Top Performers*, Product Development Institute, Inc., Ancaster, www.prod-dev.com.

Cooper, R.G. and Kleinschmidt, E.J. (1996) Winning businesses in product development: critical success factors. *Research-Technology Management*, **39** (4), 18–29.

Griffin, A. (1997) Drivers of NPD Success: The 1997 PDMA Report, Product Development & Management Association, Chicago.

Kleinschmidt, E.J., de Brentani, U., and Salome, S. (2007) Performance of global new product development programs: a resource-based view. *Journal of Product Innovation Management*, **24**, 419–441.

Levinthal, D.A. and March, J.G. (1993) The myopia of learning. *Strategic Management Journal*, **14** (Special Issue), 95.

March, J.G. (1991) Exploration and exploitation in organizational learning. *Organization Science*, **2** (1), 71–87.

Teece, D.J., Pisano, G., and Shuen, A. (1997) Dynamic capabilities and strategic management. *Strategic Management Journal*, **18** (7), 509–533.

The Economist (2008) Innovation in America: A global strorm? **389** (8607), 73–74 (Canadian edition).

front end of innovation

Peter Koen and Heidi Bertels

Definitions – What is the Front End and Why is it Important?

Innovation in large firms may be divided into three parts: front end, new-product development (NPD), and commercialization. Most research activities have focused on the NPD part (see Brown and Eisenhardt (1995) and Montoya-Weiss and Calantone (1994) for reviews), which requires a well-practiced routine and which has become regimented through the use of Stage-Gate™ (Cooper, 2001), total quality management, and six-sigma programs. The front end is typically thought of as consisting of the first three sequential stages of the Stage-Gate™ process with the remaining stages focusing on the development process. The first of these three stages is discovery, where prework is done to understand the market and generate ideas. The second stage, scoping, consists of rudimentary assessments of the market; and evaluation of the technical merits of the project. And the third stage, building a business case, focuses on more detailed investigation leading to a business case (*see also* THE STAGE-GATE IDEA TO LAUNCH SYSTEM).

More recently, Koen *et al.* (2002), Khurana and Rosenthal (1998), Langerak, Hultink, and Robben (2004), and Bertels, Koen, and Kleinschmidt (2008) have taken a more holistic view of the front end (Figure 1), which includes both organizational resources and front-end routines. Organizational resources consist of vision, strategy, resources, and climate that set the stage for the front-end routines; these, in turn, are opportunity recognition, idea generation, and concept definition. This module adds to Cooper's (2001) traditional Stage-Gate™ model by taking into account the organizational resources that lay the foundation for the front-end routines and by emphasizing the iterative nature of the front end where new learning in one of the routines can affect the activities in the others.

The front end is important because the product-development process is path dependent. This means that choices made in the front end lead to options as well as limitations regarding which products a company will ultimately be able to develop. The front end is also fundamentally different from the development portion of the innovation process (Table 1). Essentially, the front end requires more expansive and divergent thinking, whereas the NPD thrives under sequential, well-practiced routines. As a result, the management process in the front end is different from the one for the development part of the innovation process.

The front end of innovation is also referred to as the *fuzzy front end* or *predevelopment* in the academic literature. The term *front end of innovation* was first used by Koen *et al.* (2002) with the purpose of replacing the more accepted and expressionistic "fuzzy front end" first coined by Reinertsen (1985). The reasoning behind the wish to abandon the term *fuzzy front end* is that the word "fuzzy" implies that the front end is mysterious, lacks accountability, and cannot be critically evaluated. The term *front end of innovation* has gained popularity over "fuzzy front end," mostly driven by a popular practitioner conference, the "Front End of Innovation," started in 2003 and held annually.

Our discussion is divided into four sections consistent with the holistic model (Figure 1). The next section provides a definition of the organizational resources and related empirical work. The following three sections define and discuss each of the front-end routines for innovation which are, respectively, (i) exploitative, (ii) explorative but leveraging the firms' established value network, and (iii) explorative, requiring a new value network (Christensen and Rosenbloom, 1995; March, 1991). Exploitative efforts concern incremental activities, whereas explorative efforts refer to more radical or breakthrough efforts.

Organizational Resources

Organizational resources define the boundary conditions for the front-end routines. Vision, which essentially is the precursor to all of the front-end activities, is distinct from strategy. Whereas strategy provides an investment road map to innovation initiatives; a company's vision describes a future state. For example, Apple Computer's vision[1] is to bring the

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Difference between the Front End of Innovation and the New-Product Development Process.

	Front end of innovation	*New-product development*
Nature of work	Experimental, often chaotic. "Eureka" moments. Can schedule work – but not invention	Disciplined and goal-oriented with a project plan
Commercialization date	Unpredictable or uncertain	High degree of certainty
Funding	Variable. In the beginning phases many projects may be "boot legged," while others will need funding to proceed	Budgeted
Revenue expectations	Often uncertain with a great deal of speculation	Predictable with increasing certainty, analysis, and documentation as the product release date gets closer
Activity	Individuals and team conducting research to minimize risk and optimize potential	Multifunction product and/or process-development team.
Measures of progress	Strengthened concepts	Milestone achievement

Reproduced from Koen *et al.*, (2002). © John Wiley & Sons Inc., 2002.

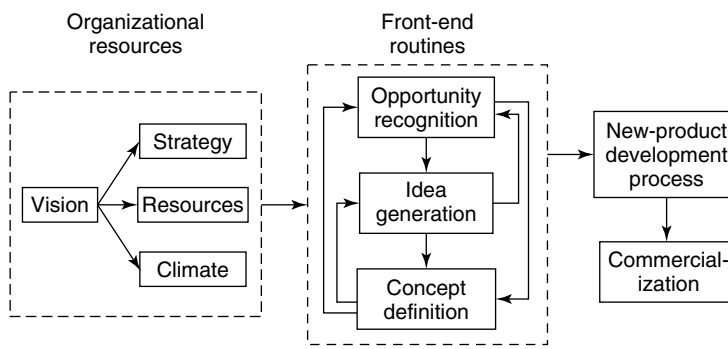


Figure 1 Holistic model of the front end.

> ... best personal computing and portable digital music and mobile experience to consumers, students, educators, businesses and government agencies through its innovative hardware, software, peripherals, services and Internet offerings.' They want to accomplish this goal through "... its unique ability to design and develop its own operating system, hardware, application software and services ..." in order to bring to the customer new products and solutions "... with superior ease-of-use, seamless integration and innovative industrial design".

In contrast, Apple's strategy defines particular products (e.g., iPod, iPhone, and Mac Computer) and product enhancements and services (e.g., iTunes and Apple stores), which would be supported within their investment horizon. Resources, in terms of both people's time and money, to support the front end represent a third key element. The final important element of the organizational resources that lay the foundation for the front end is the organizational climate.

The overall importance of vision and strategy to the front end is subject to controversy. Nonak

and Takeuchi (1995) assert that innovative ideas occur to individuals – not organizations; hence vision and strategy may be too confining and/or are relatively independent of innovative ideas rising from the bottom up. Burgelman (1983, p. 241) suggests in his study of internal corporate venturing that the "... autonomous strategic initiatives of individuals at the operational level of the organization ..." are critical for "... initiatives that fall outside the current concept of corporate strategy." In contrast, Khurana and Rosenthal (1998) hypothesized the importance of a strong linkage between organizational resources and front-end routines based on their case studies. Empirical studies by Langerak, Hultink, and Robben (2004) based on the data of 126 companies in the Netherlands, and by Bertels, Koen, and Kleinschmidt (2008) based on the data of 238 companies, mostly from the United States, found strong support that vision and strategy are significantly correlated with success in the front end.

A number of studies have confirmed the importance of climate to front-end performance. Zien and Buckler (1997), in a study of 12 leading innovative companies such as Sony, Hewlett-Packard, and Toshiba, found that an innovative climate was a consistent theme across the companies that were studied. Amabile, Hadley, and Kramer (2002), in their work on creativity, also found that a climate that allows people the time to focus on a single activity for a significant part of their work day is essential to the front end. Bertels, Koen, and Kleinschmidt (2008) hypothesized and found support for a correlation between climate and resource availability in both exploitative and explorative projects.

Front-end Routines

The three principal routines which occur in the front end are opportunity recognition, idea generation, and concept definition. Kirzner (1997) defined opportunity recognition as the recognition of an "imprecisely-defined market need or underemployed resources or capabilities." We specifically define *opportunity* as a business or technology gap that a business realizes, by design or accident, that exists between the current situation and an envisioned future to capture competitive advantage by responding to a threat, solving a problem, or ameliorating a difficulty. We define an *idea* as the most embryonic form of a new product or service. It often entails a high-level view of the envisioned solution needed to solve the problem identified by the opportunity. Finally, we define a *concept* as having a well-defined form including both a written and visual description, which includes its primary features and customer benefits combined with a broad understanding of the technology needed. The last stage can be equated with the completion of a business plan. We prefer the use of the term *concept* because a completed concept, with written and pictorial descriptions, allows the new product or service to enter the more routine part of the development process. We believe that these more precise definitions help overcome the potential confusion between opportunity recognition where the company identifies an unmet customer need and idea generation where specific ideas, which potentially respond to the opportunity are generated.

Kraft's development of Lay's light potato chip can be used as an example to better explain this terminology (Koen *et al.*, 2001). Kraft initially identified an opportunity to develop a low-fat product due to significant and high consumer interest in this area. On the basis of the recognition of this opportunity, they identified several methods for delivering nonfat potato chips using candidate molecules that could provide the same flavor, but would not be absorbed into the body. Several candidate nonfat molecules were developed resulting in Olestra, which is the ingredient used in the light potato chip. Olestra preserves the flavor of the potato chip without the accompanying fat content.

The holistic model (Figure 1) is not the same as the sequential Stage-Gate™. Ostensibly, each of the front-end routines can be mapped onto the stages of the traditional Stage-Gate™ model, that is, opportunity recognition versus discovery (Stage 1), idea generation versus scoping (Stage 2), and concept definition versus business-case development (Stage 3). We posit, however, that the front-end routines occur in each of the stages in the Stage-Gate™ model, that is, opportunity

recognition takes place – albeit to a different extent – in discovery, scoping, and business-case development. In Stage 1 there is a greater emphasis on opportunity recognition while in Stage 3 there is a greater emphasis on concept definition. We believe that this is an important difference between both models since the holistic model allows for iteration between the three routines in all stages. Iteration is an essential part of the front end. For example, working on concept definition may identify several risks and costs not initially identified, which can lead to several new ideas, which can then in turn lead to several new concepts.

Front-end Routines in Exploitation Projects

Most empirical studies on the front end have examined exploitative projects. Cooper's (2001) research, while not specifically focused on the front end, determined best-practice factors based on studies of large samples of successful and unsuccessful exploitation products. Out of the nine success factors Cooper and colleagues identified for NPD performance, three are applicable to the front end: product superiority, product definition, and up-front activities. Product superiority concerns relative product performance, that is, the benefits, quality, reduced costs, superiority, and problem-solving ability compared to competitive products. "Product definition" reflects the "whats" and "for whoms" of the product concept prior to product development, and the third success factor, "up-front activities," refers to the initial screening, preliminary market assessment, and financial analysis of the project. Similar results were found by Song and Parry (1996) who looked at new-product introductions in Japanese high-technology firms.

There have been six empirical studies that specifically look at the front end. The first four focused on exploitation projects. Bacon *et al.* (1994) indicated that the front end will lead to a well-understood product definition, which, in turn, is linked to overall product-development success. Moenaert *et al.* (1995) investigated the integration of marketing and research and development (R&D) activities and how information exchange between marketing and R&D affects success. Khurana and Rosenthal (1998) published the first comprehensive study of front-end activities on the basis of case studies of 10 incremental and 2 radical projects. They discovered that successful organizations follow a holistic approach to the front end. Furthermore, Khurana and Rosenthal (1998) classified front-end success factors into foundation and project-specific elements. Foundation elements refer to having a clear product strategy, a well-planned portfolio of new products, and an organizational structure that facilitates product development via ongoing communication and cross-functional sharing of responsibilities. The project-specific elements help clarify the product concept, evaluate it through executive reviews, and develop plans, schedules, and estimates of the project's resource requirements.

Langerak, Hultink, and Robben (2004), in their study of 126 firms in the Netherlands, developed a mediated model of market orientation on new-product performance. The mediators were strategic planning, idea generation, idea screening, and business analysis. Market orientation was measured using a 17-item construct consisting of three subscales for the company's customer orientation (i.e., understanding customer needs), competitor orientation (i.e., understanding the competitors), and interfunctional coordination (i.e., the degree to which the company's functional departments satisfy their customers). The authors found a strong correlation between market orientation and strategic planning, idea generation and idea screening – but not to business analysis. Strategic planning and idea generation were the only mediators correlated with new-product performance. The importance of a having a market-oriented focus is congruent with the work of Cooper (2001); but the lack of significant correlation between business analysis and new-product performance is not. Perhaps the explanation for this difference can be attributed to the kind of projects that were surveyed. Langerak, Hultink, and Robben (2004, p. 300) requested the survey respondents to report on "... a new product.... new to the firm ..." but where the company was "... familiar to the market." These types of projects are often classified as exploratory since they involve some degree of newness (on the product

level) and hence require learning. Cooper's (2001) study was mostly confined to exploitation projects (i.e., cost reductions, improvements to existing products, and repositioning, etc.).

Verworn, Herstatt, and Nagahira (2008) evaluated the front end of 497 projects from Japanese companies. They developed a structural equation model relating three front end factors, that is, "reduction of market uncertainty," "reduction of technical uncertainty," and "intensity of initial planning" to two NPD success factors at the project level (efficiency and effectiveness). They found a significant correlation between reduction of market uncertainty and effectiveness and reduction of technical uncertainty and both efficiency and effectiveness. Intensity of planning was significantly correlated with reduction of market and technical uncertainty as well as with efficiency. Hence, Verworn, Herstatt, and Nagahira (2008) found that the intensity of planning contributes in indirect ways to NPD success (i.e., effectiveness). The significance of a direct relationship between intensity of planning and effectiveness was not tested.

Bertels, Koen, and Kleinschmidt (2008), in their study of 238 companies, tested mediated models of predevelopment performance. The effect of organizational resources on predevelopment performance was hypothesized to be mediated by front-end routines. Predevelopment performance consisted of a six-item construct, which measured the relative competitive advantage and risk profile of the business units' front-end activities. In contrast to the earlier studies, Bertels, Koen, and Kleinschmidt (2008) developed separate mediated models for exploitation and exploration. For exploitation, they found that both opportunity recognition and idea generation were correlated with predevelopment performance, whereas concept definition was not. These results are in line with previously discussed studies (Langerak, Hultink, and Robben, 2004), which show no significant correlation between business-planning efforts and success. Business planning, based on Cooper's (2001) extensive studies, has long been considered an important contributor to overall success of the front-end activities. However, the last three studies provide empirical support for the alternative hypothesis of no significant relation. A similar debate is occurring in the entrepreneurship literature. A recent meta-analysis (Brinckmann, Grichnik, and Kapsa, in press) of 47 studies found that the business-planning process as opposed to the actual business plan had a higher correlation with success. However, no large firms were included in the meta-analysis.

Front-end Routines in Exploration Projects

Front-end routines well suited to exploration projects are less well established. The effectiveness of certain routines is dependent on whether the projects reinforce the company's current value network or require a new one. A *value network* is defined as "the context within which the firm identifies and responds to customers' needs, procures inputs and reacts to competitors" (Christensen and Rosenbloom, 1995, p. 234). It includes "its upstream suppliers; its downstream customers, retailers, and distributors; and its partners and ancillary industry players" (Christensen, Anthony, and Roth, 2004, p. 63). It is the organizational reflection of how the firm is networked to compete within a business model and encompasses the network of commitments and relationships that need to be maintained. The work of Christensen and colleagues on disruptive innovation explains why incumbent firms have difficulty developing products for new value networks.

Front-end routines for exploration projects in the existing value network. The discovery of stents is an example of an innovation that was aligned with the value network that existing companies in interventional medicine were already serving. Stents are used in a procedure for noninvasively expanding restricted cardiac arteries called *angioplasty*, which allowed companies to rapidly expand their market by solving one of the major problems associated with this procedure. Before the discovery of stents, angioplasty procedures to open obstructed cardiac arteries often could not prevent the spontaneous closure of the arteries resulting in repeat procedures. This limited the usefulness of the procedure. Stents solved this problem by preventing the artery from closing. Effective front-end routines

to develop stents are essentially the same as those discussed for exploitation projects. The company would first need to recognize that the opportunity for interventional medicine would be significantly enhanced if they could prevent spontaneous closure of the arteries. Several ideas and concepts can then be developed. The only difference from exploitation, however, is that these ideas and concepts require radically new technologies (i.e., stents).

Large firms have adopted a technology Stage-Gate™ (TSG) process (Ajamin and Koen, 2002; Schwartz, Yu, and Modlin, 2004) as modification of the traditional Stage-Gate™ process when dealing with high uncertainty radical technologies. The objective of the TSG process is to build confidence in a technology to a point where the feasibility is demonstrated and a product-development program can be initiated. The TSG lies within the front end since traditional development efforts (i.e., past Gate 3) should not be started when there is significant technical uncertainty. The TSG process holds the team accountable only to the next gate since the outcomes of the technology efforts cannot be predicted a priori. The TSG process embraces the concepts embodied in real options theory (McGrath and MacMillan, 2000) where a limited investment is made at each stage. A go/no go decision is made after each stage, depending on overall risk and value assessment based on the information collected during the previous stage. For the most part, the studies carried out in understanding the front end for exploitation projects can be applied to the routines in existing value networks.

Front-end routines for exploration projects in a new value network. Corning's optical-fiber program, General Electric's development of computerized axial tomography, Motorola's development of cellular phones, and Searle's development of NutraSweet (Lynn, Morone, and Paulson, 1996) created entirely new markets which required new value networks. Lynn, Morone, and Paulson (1996) found that the development of such products did not follow the traditional serial Stage-Gate™ routines employed for exploitation products, but followed a learning strategy, which they called the *probe and learn process*. These products required an iterative process of development and testing where early versions of the product were developed, tested in the market, tweaked, tested, and adjusted. These market probes allowed for emerging insight into what markets to pursue and what features and benefits provided value to the customer.

O'Connor and DeMartino (2006), on the basis of in-depth study of 12 radical innovation projects, identified 3 radical innovation competencies: discovery, incubation, and acceleration (*see* RADICAL INNOVATION). In the discovery stage, ideas are formulated and screened against development costs, market potential, and overall strategic value to the company. In the next stage, incubation, initial prototypes are created, which are suitable for sales and testing (this is comparable to Lynn, Morone, and Paulson (1996)'s probe and learn process). The third, or acceleration stage, requires the new product to begin to scale up and create an actual business. We have added a fourth stage, institutionalization, which considers the integration of the new opportunity with the existing business units or the development of a new business unit reporting to the traditional hierarchy of the company (Garvin and Levesque, 2005). We have developed Figure 2 to illustrate these four stages. The first three stages, while described in a sequential nature, are not sequential in practice. There is considerable learning from the second stage and third stage back into the discovery or first stage. In essence the second and third stages should also be considered as part of the front end since the learning causes new ideas to be discovered and tried out in the market place. An alternate and, perhaps, more encompassing model for

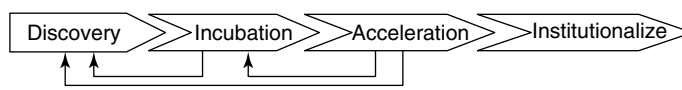


Figure 2 Front end for new value network exploration projects, which follow a probe and learn process. (adapted from O'Connor *et al.*, 2008).

radical innovation has been proposed by Reid and de Brentani (2004). This model includes the environment, the individual, and the organization. Unfortunately, this model assumes that the different stages are of a sequential nature.

Garvin and Levesque (2005) describe an emerging business opportunity (EBO) unit being used at IBM to develop opportunities that require new value networks which they describe as the "... white spaces between established businesses" (Garvin and Levesque, 2005, p. 9). The EBO unit is separated from mainstream with its initial efforts focused on understanding unmet customer needs in the marketplace. IBM makes extensive use of ethnographic methodologies (Rosenthal and Capper, 2006) whereby the development team observes the customer in his or her habitat to obtain deep insights around how the consumer experiences a product, improvises a solution, or deals with the environmental factors.

In the beginning, the strategy for a new EBO is an iterative process. People at IBM said that "sometimes it would take a year to a year and half to get the strategy we were happy with. It would change three or four times" (Garvin and Levesque, 2005, p. 10). Different from the more structured Stage-Gate™ process, review meetings for EBOs, which are held at monthly intervals, are focused on strategic clarity and understanding significant unmet customer needs. Financial proformas are used to understand the key assumptions of the project as opposed to evaluating the project on the basis of established financial hurdle rates. Project initiation and selection in the EBO structure follow the "probe and learn approach" where ultimate project approval is the result of many iterations guided by monthly senior-management dialogues. The EBO system, initiated at IBM in September 2000, has proved to be quite successful. As of spring 2003, the EBO unit had created two new business units that began generating over $1 billion in annual sales and a number of others with annual sales exceeding $100 million.

In addition to ethnographic approaches, some companies use a lead-user methodology to find new breakthrough ideas (von Hippel, 2005) (*see* LEAD USERS). Lead users, or the first inventors of a product, across seven different industries, were found to be responsible for the development of a minimum of 11% (electronic assembly) to a maximum of 100% (scientific instruments and semiconductor equipment) of breakthrough inventions. Lead users are different from other users because they are at the leading edge of an identified trend in terms of new product and process needs and because they expect to obtain a relatively high net benefit from solutions to their own need. A careful assessment of the lead-user approach at 3M (Lilien *et al.*, 2002) showed that this approach led to seven entirely new major product lines, which created $146 million (mean) in new-product sales over a five-year period as compared to the traditional approach for generating new-product lines, which created $62 million (mean) in new-product sales.

Conclusions

Compared to the work done on the product-development part of the innovation process, research in the front end remains relatively sparse. Nevertheless, organizational resources and front-end routines required for exploitation projects and for exploration projects that fit in firms' established value networks are thought to be understood. In contrast, findings related to the organizational resources and front-end routines required for new value network exploration projects are just beginning to emerge. What is apparent is that these new routines require a more iterative or learning approach. However, most companies have difficulty adopting this new approach since the skills and competencies involved are so different from those used in exploitation projects (March, 1991).

Endnotes

[1] From Apple's December 29, 2008 10K Report

Bibliography

Ajamin, G.M. and Koen, P.A. (2002) Technology Stage-Gate™: a structured process for managing high-risk new technology projects, in *The PDMA Toolbook One for New Product Development* (eds P. Belliveau, A. Griffin, and S.M. Somermeyer), John Wiley & Sons, Inc., New York, pp. 267–295.

Amabile, T.M., Hadley, C.N., and Kramer, S.J. (2002) Creativity under the gun. *Harvard Business Review*, **80** (8), 52–61.

Bacon, G., Beckman, S., Mowery, D., and Wilson, E. (1994) Managing product definition in high-technology industries: a pilot study. *California Management Review*, **36** (3), 32–56.

Bertels, H.M.J., Koen, P.A., and Kleinschmidt, E.J. (2008) Lessons from 238 companies: senior management, exploitation, exploration, and corporate entrepreneurship performance. *Frontiers of Entrepreneurship Research*, 577–593.

Brinckmann, J., Grichnik, D., and Kapsa, D. Should entrepreneurs plan or just storm the castle? A meta-analysis on contextual factors impacting the business planning-performance relationship in small firms. *Journal of Business Venturing*. In press.

Brown, S.L. and Eisenhardt, K.M. (1995) Product development: past research, present findings, and future directions. *Academy of Management Review*, **20** (2), 343–378.

Burgelman, R.A. (1983) Corporate entrepreneurship and strategic management: insights from a process study. *Management Science*, **29** (12), 1349–1364.

Christensen, C.M., Anthony, S.D., and Roth, E.A. (2004) *Seeing What's Next: Using Theories of Innovation to Predict Industry Change*, Harvard Business School Publishing Corporation, Boston.

Christensen, C.M. and Rosenbloom, R.S. (1995) Explaining the attacker's advantage: technological paradigms, organizational dynamics, and the value network. *Research Policy*, **24**, 233–257.

Cooper, R.G. (2001) *Winning at New Products: Accelerating the Process from Idea to Launch*, 3rd edn, Perseus Publishing, Cambridge.

Garvin, D.A. and Levesque, L.C. (2005) Emerging Business Opportunities at IBM, Harvard Business School Cases.

von Hippel, E. (2005) *Democratizing Innovation*, The MIT Press, Cambridge.

Khurana, A. and Rosenthal, S.R. (1998) Towards holistic "front ends" in new product development. *Journal of Product Innovation Management*, **15** (1), 57–75.

Kirzner, I.M. (1997) Entrepreneurial discovery and the competitive market process: an Austrian approach. *Journal of Economic Literature*, **35** (1), 60–85.

Koen, P.A., Ajamian, G.M., Boyce, S. *et al*. (2002) Fuzzy front end: effective methods, tools, and techniques, in *The PDMA Toolbook for New Product Development* (eds P. Belliveau, A. Griffin, and S.M. Somermeyer), John Wiley & Sons, Inc., New York, pp. 5–35.

Koen, P., Ajamian, G., Burkart, R. *et al*. (2001) Providing clarity and a common language to the "fuzzy front end". *Research-Technology Management*, **44** (2), 46–55.

Langerak, F., Hultink, E.J., and Robben, H.S.J. (2004) The role of predevelopment activities in the relationship between market orientation and performance. *R&D Management*, **34** (3), 295–309.

Lilien, G.L., Morrison, P.D., Searls, K. *et al*. (2002) Performance assessment of the lead user idea-generation process for new product development. *Management Science*, **48** (8), 1042–1059.

Lynn, G.S., Morone, J.G., and Paulson, A.S. (1996) Marketing and discontinuous innovation: the probe and learn process. *California Management Review*, **38** (3), 8–37.

March, J.G. (1991) Exploration and exploitation in organizational learning. *Organization Science*, **2** (1), 71–87.

McGrath, R.G. and MacMillan, I.C. (2000) Assessing technology projects using real options reasoning. *Research-Technology Management*, **43** (4), 35–49.

Moenaert, R.K., De Meyer, A., Souder, W.E., and Deschoolmeester, D. (1995) R&D/marketing communication during the fuzzy front-end. *IEEE Transactions on Engineering Management*, **42** (3), 243–258.

Montoya-Weiss, M.M. and Calantone, R. (1994) Determinants of new product performance: a review and meta-analysis. *Journal of Product Innovation Management*, **11** (5), 397–417.

Nonaka, I. and Takeuchi, H. (1995) *The Knowledge Creating Company*, Oxford University Press, Oxford.

O'Connor, G.C. and DeMartino, R. (2006) Organizing for radical innovation: an exploratory study of the structural aspects of RI management systems in large established firms. *Journal of Product Innovation Management*, **23** (6), 475–497.

O'Connor, G.C., Leifer, R., Paulson, A. S. and Peters, L. S. (2008). Grabbing Lightning: Building a Capability for Breakthrough Innovation (1 ed.). San Francisco, CA: Jossey-Bass.

Reid, S.E. and de Brentani, U. (2004) The fuzzy front end of new product development for discontinuous innovations: a theoretical model. *Journal of Product Innovation Management*, **21** (3), 170–184.

Reinertsen, D.G. (1985) Blitzkrieg product development: cut development time in half. *Electronic Business*.

Rosenthal, S.R. and Capper, M. (2006) Ethnographies in the front end: designing for enhanced customer experiences. *Journal of Product Innovation Management*, **23**, 215–237.

Schwartz, K.J., Yu, E.K., and Modlin, D.N. (2004) Decision support tools for effective technology commercialization, in *The PDMA Toolbook Two for New Product Development* (eds P. Belliveau, A. Griffin, and S.M. Somermeyer), John Wiley & Sons, Inc., Hoboken, pp. 435–460.

Song, M.X. and Parry, M.E. (1996) What separates Japanese new product winners from losers? *Journal of Product Innovation Management*, **13**, 422–439.

Verworn, B., Herstatt, C., and Nagahira, A. (2008) The fuzzy front end of Japanese new product development projects: impact on success and differences between incremental and radical projects. *R&D Management*, **38** (1), 1–19.

Zien, K. and Buckler, S. (1997) Dreams to market: crafting a culture of innovation. *Journal of Product Innovation Management*, **14**, 274–287.

quality function deployment (QFD)

John R. Hauser, Abbie Griffin, Robert L. Klein, Gerald M. Katz, and Steven P. Gaskin

Definition

Quality function deployment (QFD) is a product-development methodology whose objective is to "deploy" the VOICE OF THE CUSTOMER (VOC) throughout the product-development process. It is most often carried out by a CROSS-FUNCTIONAL TEAM that creates and populates a series of one or more matrices, the first and most common of which is referred to as the *House of Quality* (HOQ). When completed, these matrices relate customer wants and needs (the VOC) to an extensive set of product features. A set of metrics is developed to measure how well any given set of product features is meeting customer needs. These metrics are then used to identify and prioritize the product's design specifications (*see* PRODUCT SPECIFICATIONS).

Description and Commentary[1]

QFD is believed to have been developed in the early 1970s at a Japanese shipbuilding firm. Its objective was to provide a systematic way of dealing with the many complexities and trade-offs inherent in all the design decisions faced by product developers. After considerable study and a number of improvements by several important Japanese academics, it later migrated to the Japanese auto industry and then to the US auto industry by the mid-1980s (Hauser and Clausing, 1988). Today, it is used in almost every type of industry and application imaginable – products and services, consumer (B2C) and commercial (B2B) applications, high tech and low tech industries, and so on.

QFD is best carried out by an active cross-functional team whose job is to complete one or more of a series of matrices that lead to a set of insights about how best to create a winning product or service and how to prioritize their research and development activities going forward. Many QFD practitioners believe that most of the learning takes place during the process of debate among the cross-functional team members in trying to reach consensus on the various entries into the matrix.

A rigorous view of QFD employs not just one matrix (the original HOQ), but four. These are sequentially related as follows:

- customer needs to performance measures
- performance measures to features or solutions
- features or solutions to parts specifications
- parts specifications to manufacturing processes.

A variation on the second of these matrices is called the *Pugh concept selection*, developed by Prof. Stewart Pugh in the United Kingdom. This variation attempts to evaluate different product concepts against the key performance measures, with the objective of incrementally moving toward an ideal concept.

There are many different styles of conducting QFD, and practically no two practitioners follow exactly the same process. A particularly lively debate has to do with the importance of completing all four matrices, with many practitioners arguing that most of the value is derived from the first matrix alone.

It is well recognized that customers choose products on the basis of how well the products fulfill their perceived needs. New-product development teams must select product features to fulfill those perceived needs.

Figure 1 illustrates a stylized HOQ completed by Puritan-Bennett, a medical-device manufacturing company, when they were redesigning a medical instrument called a *spirometer* (a device that measures lung capacity). The HOQ begins with a formal study of perceived customer needs called the "*Voice of the Customer*." The VOC identifies customer needs such as "the product is easy for the physician to hold while taking measurements on a patient." In the Puritan-Bennett example, interviews with physicians, technicians, nurses, and patients identified 25 strategic customer needs. These are listed on the left-hand side of the matrix.

However, all needs are not equally important. The customers would much prefer that some needs be fulfilled even if it means that other

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

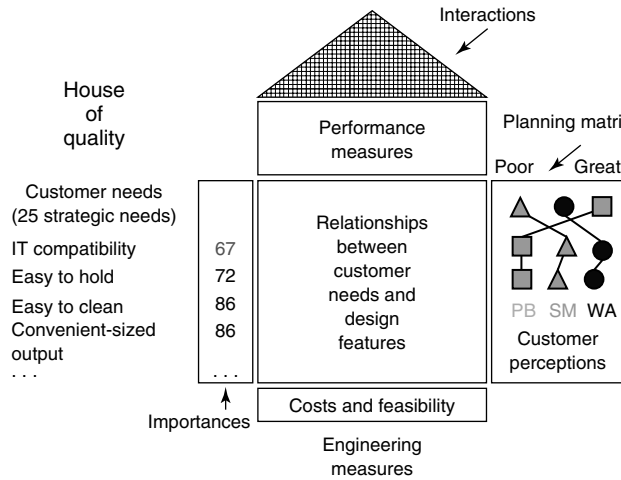


Figure 1 Illustrative house of quality.

needs are not fulfilled as well. The VOC, in addition to defining customer needs, also measures the importance of each need. In Figure 1 these measures of importance are listed to the right of the list of customer needs. In addition, the HOQ lists on the right how each existing product (shown for three manufacturers that we label PB, SM, and WA) fulfills the customer needs. For example, WA does extremely well on the important customer need of affordability, while PB does poorly. The diagram shows that affordability is a high-priority need.

The remainder of the HOQ is reasonably self-explanatory. The product-development team lists performance measures at the top of the house. For spirometers, a performance measure might be the weight, the diameter, or the number of minutes required to clean it.

The team then considers each high-priority customer need, such as "information technology (IT) compatibility," and, in the center of the house, indicates how each performance measure affects the fulfillment of that need. For example, a small diameter may make it easier for someone to hold (positive relationship), but take longer for someone to clean (negative relationship). Some teams indicate the relationships with symbols (●, ◉, ○), others with numbers (±9, ±3, ±1), and still others with simply + or −. The HOQ is a guide; any quantification must be tempered with qualitative judgment.

The triangular roof of the HOQ holds interactions between performance measures. For example, it is likely that a larger diameter is positively correlated with a higher weight product, and thus a "somewhat related" relationship would be noted at the intersection box between them in the roof. At the bottom of the HOQ are costs, technical feasibility, benchmarks, and, possibly, metrics (engineering measures) that are performance goals for this project.

Early applications use large numbers (hundreds) of customer needs and design features, and building the HOQ with that many needs becomes an arduous task. QFD has become somewhat controversial in many organizations owing to the time, effort, and (some would argue) tedium that it involves. Fortunately, excellent software and a number of interesting shortcuts are now available to combat these problems. For example, the HOQ has evolved through the use of "Turbo HOQ" to a representation that is more "just in time." When the team considers a customer need, it fills in the design features to which that need is linked.

Though QFD can involve a good deal of effort, there are a number of key benefits to be derived from its use:

- It allows teams to prioritize their development activities in a systematic, analytical way that puts the customer first, as opposed

to through a political free-for-all that relies on which customer, salesman, or officer can shout the loudest and exert the most power.
- It takes advantage of cross-functionality in an orderly, truly participative way, enlisting the support of all major functions within the organization toward a common view.
- It provides an "audit trail" that reminds people, both new and old to the project, as to why certain decisions were made in the past.
- It often results in a prioritization that is highly unexpected and different from the conventional wisdom held by the company and many of the participants before engaging in QFD, thus stretching the team's thinking as to which activities are most critical toward creating a winning product or service.

Having noted the benefits, let us turn to some examples showing why and how QFD can be helpful in the product-development process.

Much of the effort involved in QFD is devoted to the relationship between customer needs and features during the QFD process. Why is it so hard to relate customer needs to engineering features? Consider a liquid dishwashing detergent (for washing dishes by hand). It is basically a chemical product. Customers buy it to clean their dishes. But what does "clean" mean to customers in the context of dishes and how do customers judge that their dishes are clean? It is unlikely that they use a magnifying glass or a scientific instrument to measure the light reflected from the dishes, although an engineer might use those instruments to test a dishwashing product's performance. It is also unlikely that the customer will read and understand the chemical ingredients. More likely, the customer will use some subjective means to determine "clean." This might mean holding dishes up to the light or it might simply mean that there is no noticeable dirt on the dishes – a minimum requirement. Tactile cues, such as running fingers along a dish to detect the presence of grease, may play a role. In addition, the customer might derive peripheral cues, such as the clarity of the water in the sink (after washing dishes) or the amount and type of bubbles that are still around when the dishes are done and ready to be rinsed. The scent of the dishwater or of the dishes might be another cue to cleanliness.

But "clean" might not be the only perceived need. The customer might care about the ease of use, the "feel" of the water while washing the dishes, the scent of the liquid, the effect of the dishwashing liquid on skin, the ease of storing the bottle (or package), whether the washing causes the dishes to deteriorate, how much is needed to wash a sink full of dishes, or whether the liquid imparts a perceived taste to the dishes. It is not enough to engineer the best cleaning liquid. The product-development team must engineer the entire customer experience. This includes the liquid itself as well as the packaging and any advertising image.

Consider another example – engineering a telephone-service center to provide great service. The development team may design the physical space, select the telephonic equipment, and design protocols. The development team may also design a personnel policy, training, and a monitoring/reward system. Consider the monitoring/reward system. One major financial service provider determined that customers did not like to wait. They instituted metrics and rewarded the service providers to minimize the number of rings before answering, avoid transferring customers, and answer the customers' questions as rapidly as possible. Unfortunately, the service representatives soon figured out how to "game" the system: answer the phone quickly, avoid a transfer even if the service provider did not know the answer, and get off the phone as quickly as possible. In fact, the metric was improved by giving incomplete and unhelpful answers.

The financial services firm responded. Service providers were now given incentives to stay on the phone until the customer got an answer. One metric was the number of minutes per hour that the service provider was on the phone. The service providers responded. They were never off the phone. Breaks were taken while the customer was on hold.

Ultimately, the financial services firm talked to the service providers and understood their needs. Service providers were not happy with gaming the system. They wanted to serve customers but felt that the metrics prevented them from doing so. In response, the firm began using more qualitative metrics based on monitoring and on customer satisfaction scores. Service providers

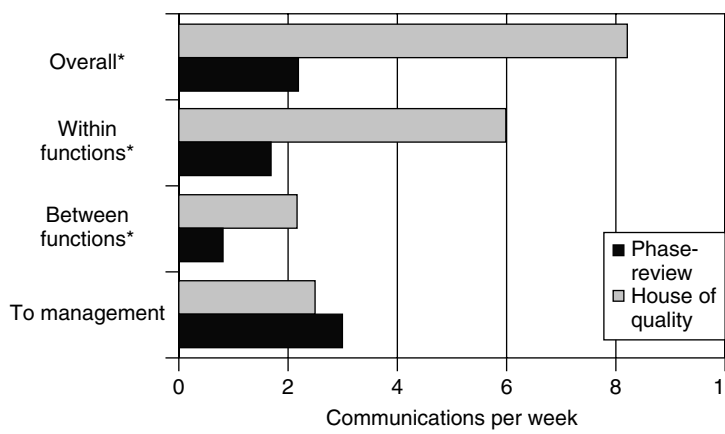
started to maximize the customers' perceived needs – the customers wanted their questions answered correctly. Efficiency was a secondary criterion.

Many other examples abound. Fulfilling customer needs is important to the success of a new product, but it is difficult to achieve. It is especially difficult in a complex product such as an automobile or a high-end copier. An automobile may take 1000 person-years to design – millions of decisions need to be made. Even an office copier may require close to 10 000 critical engineering decisions. Effective product developers want every one of those decisions to be focused on the customer.

Since customer needs drive the entire QFD process, it is critical that the list of customer needs be complete; otherwise, important factors in product design may be omitted. The *Voice of the Customer* is a term used in business to describe the process of capturing a customer's requirements (*see* VOICE OF THE CUSTOMER). The VOC is a product-development technique that produces a detailed set of customer wants and needs which are organized into a hierarchical structure, and then prioritized in terms of relative importance and satisfaction with current alternatives. It involves both qualitative and quantitative market research among current and potential customers. Sometimes, practitioners are tempted to gather the VOC in a very cursory way for reasons of time, budget, or lack of interest in what their customers have to say. This is rarely a good strategy – the quality of the output depends upon the quality of the input. Gathering a good VOC takes time, and requires knowledge of sample design, recruiting of research respondents, good interviewing techniques, the ability to translate interview output into a useful set of needs, and so on. For practitioners new to the VOC process, it is good practice to obtain some training from experienced VOC practitioners before attempting to conduct this research, or to outsource this task to them (for more information *see* VOICE OF THE CUSTOMER).

Another major benefit of QFD is that it improves communication among members of the product-development team. Here is an example. In the early days of QFD, MIT Sloan researchers studied its implementation at Ford Motor Company (Griffin and Hauser, 1992). Two teams were chosen. Each was working on a similar, but different, component of a new automobile and each team was otherwise similar in terms of skills and team members. Both teams reported to the same managers. One team used the QFD/HOQ; the other team used Ford's standard phase-review process. Over the course of the product-development cycle, the MIT Sloan researchers measured the amount and type of communication among team members. The results are given in Figure 2.

Overall, there was significantly more communication among members of the QFD team.



*Significant at the 0.05 level.

Figure 2 Communication at Ford.

Deeper analysis revealed that QFD focused its communications within the team – both within functions and between functions in these interfunctional teams. The only type of communication reduced by the HOQ was that between the QFD team and management. Even deeper analysis showed that the phase-review team engaged in significantly more "up-over-down" communication. For example, an engineer at Ford might communicate a design change to his/her manager who would communicate that design change to a manager at a supplier who would, in turn, communicate the design change to an engineer at the supplier. In the HOQ team, the Ford engineer communicated directly with the engineer at the supplier.

Figure 2 is just one example of the success of the various means to enhance communication among product-development team members concerning the relationship between perceived customer needs and engineering design features. Figure 2 is based on QFD, but there are many other ways to effect communication during the product-development process. For example, Wind *et al*. (1989) provide an excellent example of how conjoint analysis was used to design "Courtyard by Marriott." They show how, in addition to the usual quantification of trade-offs, conjoint analysis was used to link the features of the hotel to perceived customer needs. Green, Krieger and Wind (2004) provide another example where conjoint analysis was used to design the EZPass (FastLane) electronic toll-payment system (Green, Krieger, and Wind, 2004).

The important lesson here is that such communication must occur if successful products are to be designed. It is far more efficient and effective to incorporate the VOC into the new-product development process proactively, and early on, through a systematic process such as QFD, than it is to redesign the product after an unsuccessful launch.

Endnotes

[1] Significant portions of this article are drawn from an MIT Sloan Courseware document by John R. Hauser, "*Notes on 'Engineering' Product Design*," MIT, Cambridge, MA 2008. John R. Hauser and MIT Sloan grant a nonexclusive right to use this material in this description of QFD. John R. Hauser and MIT Sloan retain a nonexclusive right to this material.

Bibliography

Akao, Y. (2004) *Quality Function Deployment: Integrating Customer Requirements into Product Design*, Productivity Press, Inc., New York.

Akao, Y. and Mizuno, S. (1994) QFD: The Customer-Driven Approach to Quality Planning & Deployment, Asian Productivity Organization, Tokyo.

Cohen, L. (1995) *Quality Function Deployment: How to Make QFD Work for You*, Addison-Wesley Publishing, Reading.

Green, P.E., Krieger, A., and Wind, J. (2004) Thirty years of conjoint analysis: reflections and prospects, in *Market Research and Modeling: Progress and Prospects, a Tribute to Paul E. Green* (eds J. Wind and P.E. Green), Kluwer Academic Publishers, New York.

Griffin, A. and Hauser, J.R. (1992) Patterns of communication among marketing, engineering and manufacturing – a comparison between two new product teams. *Management Science*, **38** (3), 360–373.

Hauser, J.R. and Clausing, D.P. (1988) The house of quality. *Harvard Business Review*, **66** (3), 63–73.

Katz, G. (2007) Quality function deployment and the house of quality, in *The PDMA Toolbook 3 for New Product Development* (eds A. Griffin and S. Somermeyer), John Wiley & Sons, Inc., Hoboken.

King, B. (1987) *Better Designs in Half the Time: Implementing QFD in America*, GOAL/QPC, Methuen.

Wind, Y., Green, P.E., Shifflet, D., and Scarbrough, M. (1989) Courtyard by Marriott: designing a hotel facility with consumer-based marketing models. *Interfaces*, **19**, 25–47.

creativity

Jacob Goldenberg and David Mazursky

Introduction

Karl Popper expressed traditional thinking about creativity when he stated, "Creativity is a divine spark that may not be dismantled and examined by use of scientific tools." It is not surprising, therefore, that until the 1970s, no thorough scientific investigation of creativity has been conducted. For many years, the study of creative thinking and methods that render this process more effective remained the domain of practitioners such as engineers or advertisers, who sought tools to enhance their ordinary problem-solving tasks. In the last 30 years, however, many creativity-enhancement methods have been proposed and many have enjoyed immense popularity. Common to most of these methods is the belief that to ignite the creative spark, all we have to do is break away from existing conceptual frameworks, suspend all judgment and criticism, and search diligently for the nonconventional. A small minority propounds a more systematic approach to creativity enhancement.

In the academic world, the pioneers who broke the "taboo" on research in this field were cognitive psychologists, who were later joined by investigators in other disciplines such as neuropsychology, artificial intelligence, engineering, education, and, more recently, marketing and management science researchers. In this way, it was long after the adoption of various creativity-enhancement methods by practitioners that academic research began to investigate the common beliefs and underlying assumptions of these methods, and construct a scientific knowledge base in this area.

Here, we focus on creativity in new-product ideation with an attempt to conceptualize what is known today. An introduction to definitions of creativity is followed by an overview of the underlying assumptions about effectively developing creative products, and contemporary directions of research on creativity. We explain two very different approaches to enhance creativity in greater detail, including the numerous variations that have been suggested and are in use. The discussion section concludes with some thoughts on future progress.

Definition

Creativity is considered the ultimate of human qualities and one of the key measures of intelligence that is manifest in many fields. Creativity is also an enigmatic phenomenon. Like intelligence, it represents a highly complex and diffuse construct revealed by the wealth of available definitions for creativity (Sternberg, 1985, 1988). Despite various differences in these definitions, *creativity* is often defined as useful novelty – "not novelty for its own sake, but novelty that can be applied and adds value" (Oldham and Cummings, 1996), and most definitions share these two elements: (i) novelty or originality and (ii) usefulness, value, or benefit (Boden, 1995; Mumford and Gustafson, 1988). Mayer (1999, p. 449), for example, notes that most authors who contributed to the Handbook of Creativity endorse the idea that creativity involves the creation of an original and useful concept. Table 1 summarizes literature (mainly outside new products) that holds that the two main components of creativity are originality and usefulness: most works, however, do not offer measures of the various dimensions.

This accepted conceptualization of creativity has also been confirmed by findings of a recent study (Goldenberg *et al.*, 2009), which used *similarity structure analysis* (*SSA*) (Guttman, 1968; Shye and Elizur, 1994), a type of multidimensional scaling (MDS) technique that provides spatial representations of the relations among variables: the more highly correlated a pair of items, the closer these two items are located in the space. The SSA analysis performed on 19 new products, revealed a three-dimensional structure (see Figure 1) that confirms the perception that creativity is a combination of originality and usefulness. It is therefore not surprising that researchers in marketing refer to and use these dimensions when studying creativity.

Creative ideas are also believed to contribute to the thinking atmosphere and the enhancement of productivity, even when they do not produce immediate profit (Weisberg, 1992). "Generating interesting designs is desirable not only because

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Dimensions of creativity.

Originality	*Usefulness*	*Source*
Novelty	Appropriateness	Amabile (1983) and Jackson and Messick (1965)
Novelty	Usefulness	Gilford (1967), Ray and Myers (1986), Mumford and Gustafson (1988), Sternberg (1988), Osche (1990), Lubart (1994), Sternberg and Lubart (1995), Oldham and Cummings (1996), Gruber and Wallace (1999), Lumsden (1999), Martindale (1999) and Unsworth (2001)
Novelty	Socially valued	Taylor, Smith, and Ghiselin (1963)
Novelty	Tenability	Stein (1953)
Originality	Value	Sobel and Rothenberg (1980) and Young (1985)
Originality	Meaningfulness	Haberland and Dacin (1992), Thorson and Zhao (1997) and Wells (1989)
Originality	Worth or usefulness	Rothenberg and Hausman (1976)
Originality	Appropriateness	Runco and Charles (1993), Amabile (1997) and Kasof (1995)
Originality	Effectiveness	Field (1968)
Uniqueness	Meaningfulness	Sethi, Park, and Smith (2001, p. 74)
Unusualness	Appropriateness	Jackson and Messick (1965)
Surprise	Effectiveness	Bruner (1992, p. 3)
Divergence	Relevance	Smith and Yang (2004)
Divergence, difference	Fruitfulness, helpfulness, constructiveness – in short, productiveness	Tellis (1998, p. 83)

these designs will often be better than previous solutions, but also because they lead to a new way of thinking about the problem" (Ulrich, 1988).

In view of the potential gains involved, we should not be surprised at the vast resources invested in the search for guidance on how to cultivate and promote creative ideation. Efforts to study and develop methods for enhancing creativity broadly reflect one of two different answers to the question "Where is creativity located?" Accordingly, approaches and methods are characterized by either a focus on individuals, or a focus on ideas.

Seeking Creativity in the Individual

In the first approach to gaining a greater understanding of how creativity works and how it can be fostered in new-product development, the individual is viewed as holding one of two major roles: either as the creative personality who conceives the original idea, or as the repository of information on the basis of which new creative ideas can be developed by others. Investigators who follow the first view have concentrated on producing a profile of the creative individual, for example, by conducting physical examinations as well as personality tests on people who were considered to be creative.

On the basis of their studies of the creative personality in the early 1970s, cognitive psychologists constructed a quantitative yardstick for measuring creativity: a creative person was someone with a large flow of ideas in a given unit of time (Marschak, Glennan, and Summers, 1967). It was thought that from the greater number of different ideas thus produced, the creative person can cull higher quality ideas that might be absent from a smaller set of ideas.

More recent studies, however, have cast doubt on this criterion by showing that a large output of ideas does not necessarily lead to creative ideas. Furthermore, the very occupation with ordinary ideas seems to hamper creativity and innovative thought. In short, no correspondence

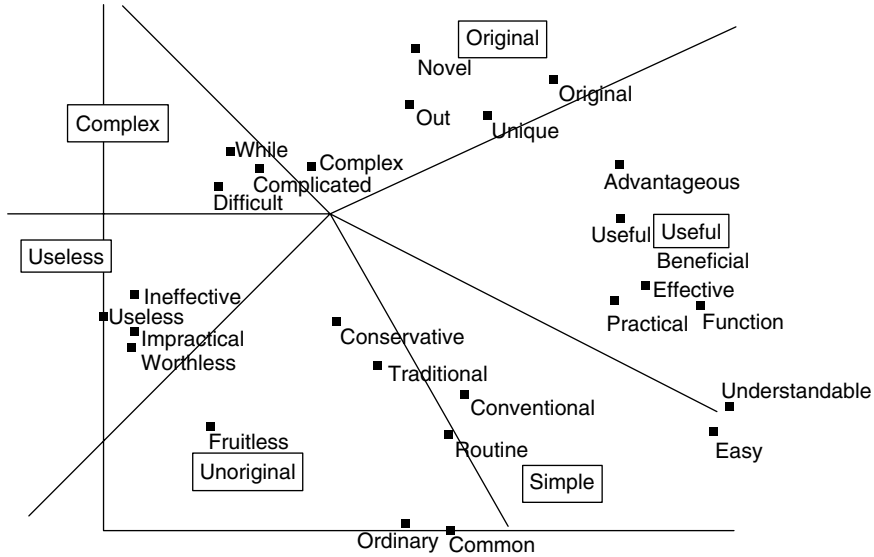


Figure 1 Dimensions of creativity.

between quantity and quality of ideas appears to exist (Connolly, Routhieaux, and Schneider, 1993; Paulus *et al.*, 1993; Stroebe, Diehl, and Abakoumkin, 1991).

The second view of the individual's role in the creative ideation process is epitomized by the market research approach, which sees the individual as the major input and source of information for the creative ideation process. Marketing theory offers many tools for market analysis and data mining designed to promote original ideas by tapping into consumers' views in the search for unfulfilled needs in the market that can be translated into creative new-product offerings. Some individuals may be more instructive than others in this context. According to Von Hippel (Urban and von Hippel, 1988; Von Hippel, 1989), valuable information may be gathered by being attentive to "LEAD USERS". Von Hippel indicates that these users serve as an "ideational elite" and possess unique information about future needs. By creating solutions to their own problems, they are frequently able to provide information for the development of new products, and to predict their success.

The ability of market research, or "the consumer's voice" (*see* VOICE OF THE-CUSTOMER), to elicit truly *original* concepts has come increasingly into question, and specific arguments have recently criticized the value of marketing research as a source of new-product ideas (Durgee, O'Connor, and Veryzer, 1996). In fact, a close look at historical data reveals that many new and surprising products were never developed from market research; Even products that have entered the "Hall of Fame" of innovation and creativity were envisioned by a diverse group of individuals, some of whom had no knowledge at all about the market for the products they developed, and some never even imagined that the market would be interested in their inventions.

These critics explain that consumers are simply not a beneficial source of information of the kind that triggers creative ideation: many desires and needs lie below the surface of consciousness, and current product users are not able to clearly articulate for nonexisting products. Griffin and Hauser (1993) also expressed doubts about consumers' ability to predict exactly which products the firm should develop, the details and features of the future blockbuster products. More generally, researchers doubt whether consumers provide reliable information beyond their personal experience or knowledge (cf., Griffin and Page, 1996), Von Hippel (1984)

summarized the problem succinctly: "Clearly, the average user of today's product has no experience with tomorrow's products . . . the average user is in a poor position to provide accurate data about future products to market researchers."

The use of market research data as a source for creative product ideas – ideas that are both useful and novel – is further undermined by the low probability of an exclusive discovery of an emergent need and the subsequent introduction of an innovation to the market (for more details see Goldenberg, Mazursky, and Solomon, 1999a). In view of the aggregate dynamics of an emerging desire or need, and its propagation in the market, Goldenberg and Efroni (2001) conclude that by eliciting information from consumers, a firm operating in a market with few competitors may discover an emerging need when the gradient of the awareness propagation process is large, that is, this discovery is less likely to lead to a truly new innovative product idea because an increasing number of competitors are simultaneously in the process of its discovery.

While market research may be imperative for understanding how products should be modified to meet market needs, it can determine the best design of a product given market preferences, predict the product's success, and ascertain the most appropriate time to launch the new product (*see also* PRODUCT TESTING). However, as a representation of market reality at a specific time, marketing data offers a poor indication of future market needs, and thus offers a poor foundation for creative product ideation.

Focusing on the Ideation Process

Look outside the box?. A common lore is that creativity exists "outside the box." The source of this famous notion seems to be the nine-dot puzzle, which may be the only problem that actually has a box, and whose solution exists beyond itself. How reliable is this belief?

Creativity is a young scientific discipline whose roots can be traced to the early 1970s, when a psychologist named Guilford was one the first researchers to conduct a scientific study of creativity. One of Guilford's most famous studies was the nine-dot puzzle, presented with its solution (Figure 2). The problem is to connect all nine dots using four lines, without lifting your pencil from the page. Today, more people know about this puzzle, but in the 1970s very few were aware of its existence despite its long history.

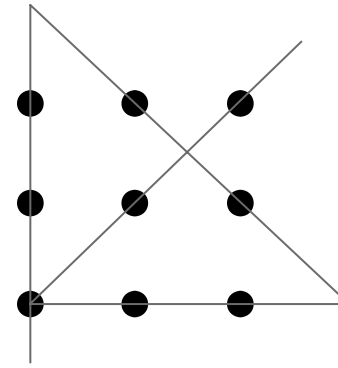


Figure 2 The nine-dot puzzle and its four-line solution.

The correct solution to this problem requires the solver to draw lines that extend *beyond* the area confined by the dots, yet would-be solvers usually confirm that their first attempts at a solution typically involve sketching lines *inside* the imaginary square. All the participants in Guilford's original study were similarly ensnared by the imaginary square and were unable to "see" the white space beyond the square's boundaries. Only 20% managed to break out of the illusory confinement and extend their lines into the white space surrounding the dots. The beautiful simplicity of the solution and the fact that 80% of the participants were mesmerized by the square while only 20% of all participants solved this puzzle correctly, led Guilford and the readers of his publications to leapfrog to the sweeping conclusion that creativity requires one to go "outside the box." Overnight it seemed that creativity gurus were preaching and teaching managers how to enhance their creative potential by thinking "outside the box."

The nine-dot puzzle became an icon for creativity and the phrase "thinking outside the box" rapidly spread in marketing, management, psychology, personal improvement. There seemed to be no end to the insights that could be offered under the banner of "thinking outside the box." Speakers, trainers, training-program

developers, organizational consultants, university professors – all had much to say about the vast benefits of outside-the-box thinking. The concept enjoyed such strong popularity and intuitive appeal that few bothered to check the facts.

In the 1970s, around the same time as Guilford conducted his experiments, another much less known study was conducted (Weisberg, 1992). This experiment used the same puzzle, but a different procedure. Participants were divided into two groups: the first group was given the same instructions as the participants in Guilford's experiment. The second group was informed that the solution requires the lines to be drawn outside the imaginary square bounded by the dot array. Surprisingly enough, only 25% of the group that received the clue solved the puzzle correctly, which is not a statistically significant difference.

Even though the solution requires participants to literally think "outside the box," participants' performance was not improved even by specific instructions to literally look for the solution *outside the box*. In other words, direct explicit instructions to think outside the box made no impact in improving solution rates.

Inside the box?. The nine-dot puzzle study illustrates the underlying flaws of this concept, even though it is frequently used (in abridged form) by many creativity lecturers. The remarkable case of the "Fosbury Flop" (Goldenberg *et al.*, 2010) has featured so frequently at creativity conferences that it has become almost as iconic a symbol of thinking outside the box as the nine-dot puzzle.

The year 1968 is seared in world memory as a phenomenal year in athletics. In high-altitude oxygen-poor Mexico City, Bob Beamon's 8.90-m jump was hailed as the greatest athletic achievement of all time. Beamon's achievement exceeded the previous world record by 55 cm, and remained unbroken for 23 years. Yet in a different corner of the stadium, an unknown athlete was responsible for one of the most dramatic and sensational revolutions that occurred in the history of sports. Introducing a radical innovative technique, Dick Fosbury won the gold medal in the high-jump event with a back-first flip that he himself had invented. This technique was contrary to the conventional high-jump technique known as the *straddle*, or the *western roll*. With a technique that had never been seen before, Fosbury revolutionized the entire sport. In less than 10 years, virtually all high jumpers had adopted his technique, and the previous high-jump technique became obsolete. This newly embraced method was named the *Fosbury Flop* after its charming and unassuming inventor.

Interestingly, the case of the Fosbury Flop often is used in conventions and features in conference speakers' materials to support the idea that revolutions originate from out-of-the-box thinking. After all, the Fosbury Flop prompted a genuine revolution in the field of high jump, using a technique that was almost the polar opposite to the dominant technique at the time. All athletes used the straddle, in which the jumper approaches the bar with his face toward the bar, and jumps and rolls over the bar with his stomach toward the bar, but Fosbury approached the bar with his side, turning his back to the bar when rolling over it. That he literally used the "opposite" technique was taken as clear evidence that Fosbury was thinking "outside the box."

The true account, however, illustrates that Dick Fosbury's revolution had nothing to do with thinking outside the box. There is no argument that Dick Fosbury's idea was creative – it was unique and it effectively solved the problem of improving jumpers' performance. But according to Dick Fosbury's own account of the Fosbury Flop development, there were no leaps in the thought process, it was the result of ordinary thinking: his new idea involved no outside the box thinking.

Fosbury's Flop technique in the high jump began as a result of his first having learned an antiquated, energy-wasting technique called the "*scissors*" at the age of 10. In the fifth grade, at the age of 11, Fosbury's physical education teacher and coach taught all the children trying out for track to jump using the classic style, the *straddle*. Fosbury, however, continued to use the scissors jump until he reached high school, mainly because he was not able to master the straddle enough to improve his performance.

In high school, the scissors were no longer accepted and, in switching to the straddle, Fosbury effectively had to learn how to jump

all over again. His competition results were far behind the other jumpers. Extremely frustrated, Fosbury asked his coach whether he could revert to the old scissors style to improve his results, and maybe boost his confidence a bit. Although he was not enthusiastic, the coach was sympathetic to the young athlete's frustrations and agreed to let him try it. So in a fateful career decision, instead of working to improve his straddle skills, Fosbury reverted to the technique he felt comfortable with, even if it was much older and less "efficient."

Fosbury decided to try his old style on his next competition. Feeling awkward yet determined, Fosbury cleared his previous best jump at 5′ 4″, but when he faced a new height, he understood that something in the technique had to be changed. With the scissors style, the jumper typically hits the bar off with his/her behind, and sometimes with the movement of the legs. To compensate, Fosbury began to try to lift his hips up higher, which also dropped his shoulders simultaneously. He continued to raise his hips until he eventually cleared 6″ and set a new personal record, and even placed fourth to score points for his team.

No one noticed what Fosbury was doing, because he was tweaking the old technique into something new, one tiny step at a time. Each attempt was only marginally different from the previous one. When Fosbury unexpectedly began to beat the other competitors, his rivals' coaches noticed that the athlete was using "something different." When they checked the rule book, they could find no evidence for anything illegal in his strange technique. Fosbury was simply applying incremental improvements. There was no radical change or leap into uncharted territory – Fosbury methodically made incremental changes, all keeping within the scissors concept, until finally he cleared the bar with his back to the bar, arching his hips over, then unarching to kick his heels over and land on his back in the pit. At the 1968 Olympics Fosbury won the gold with his Flop.

While many creativity speakers tell this story to show that Fosbury was thinking "outside of the *straddle* box," when the facts are revealed, it can be seen that this is not the case: Fosbury was, in fact, thinking "inside the *Scissors*-jump box."

A process like No other?. Closely related to the diametric approaches of out-of-the-box and inside-the-box thinking, is the division among scholars and practitioners concerning the nature of the creative process itself. Many scholars state that the creative process is a unique occurrence, one that is qualitatively different from ordinary, day-to-day conventional thinking (Guilford, 1950; Koestler, 1964; Mackinnon, 1970; Wallas, 1926).

According to this line of thought, creativity is conceptualized as involving a leap that cannot be adequately formulated, analyzed or reconstructed, a task that requires "complete freedom of symbolic expression" that "fosters the openness, the playful and spontaneous juggling of percepts, concepts, and meanings" (Rogers, 1959) it approaches "the unknown, the mysterious, the puzzling" (Maslow, 1959); it is "often obscure, unknown, unperceived even by the person himself" (Anderson, 1959). This sense of a leap is manifested by numerous descriptions and personal testimonies of creativity as emerging from "thin air," or even from an apparently complete "void." Sinnott argues that "it is common for a new idea to arise almost spontaneously in the mind, often seemingly out of nothing and at a time when a person may be thinking of something quite different" (Sinnott and Anderson, 1959). Helmholtz stated, "[Creative ideas] often enough crept quietly into my thinking . . . they were simply there . . . But in other cases they arrived suddenly, without any effort on my part, like an inspiration . . . Often they were there in the morning when I awoke" (Woodworth, 1938).

Poincaré (1952) describes his work on a mathematical problem in the same vein and in a casual manner: "One day, as I was crossing the street, the solution of the difficulty which had brought me to a standstill came to me all at once." Mozart likewise accounts: "When I am, as it were, completely myself, entirely alone, and of good cheer – say, traveling in a carriage, or walking after a good meal, or during the night when I cannot sleep; it is on such occasions that my ideas flow best and most abundantly" (Mozart, 1952, p. 34).

Guilford has shown that most of the aptitude factors identifiable as belonging in the category of creativity may be classified as a group

of divergent-thinking abilities. These abilities (evaluated by such tests as the "unusual uses"), in contrast to regular convergent-thinking abilities, emphasize searching activities with freedom to go in different directions, if not a necessity to do so to achieve excellent performance (Guilford, 1959, p. 161). The creative process is inherently fuzzy: fuzziness works where contradictions begin and opposites fuse in a paradoxical ambiguity (Kosko, 1992).

Nonetheless, creativity researchers remain divided over the extent to which the creative process itself has a distinctive nature (Maimon and Horowitz, 1999), and whether creativity is an extraordinary event of thinking or an ordinary problem-solving process. Among others, scholars such as Perkins (1981, 1988) and Weisberg (1986) suggested that creativity does not necessarily require a qualitative leap or a creative spark: it is the outcome of ordinary or everyday thinking, from which it is only quantitatively different and does not necessarily require a qualitative leap or a creative spark. Weisberg (1986, p. 69) summarizes this approach as follows: "Creative thinking is not an extraordinary form of thinking. Creative thinking becomes extraordinary because of what the thinker produces, not because of the *way* in which the thinker produces it."

Methods to Enhance Creativity: A Review

Creativity-enhancement methods are numerous and diversified. The methods that have been developed, refined, and used by individuals and firms to enhance creativity reflect the underlying assumptions concerning the uniqueness or mundane nature of the creative spark, and the need to search for creative ideas either inside or outside the box.

Smith (1998) provides an analysis of 172 idea-generation techniques used by organizations and creative consultants, which belong to two broad categories of methods: unstructured methods designed to foster the unique, genius-like spark that comes from out-of-the-box thinking, and structured methods designed to impose order on the search for creative ideas (*see* IDEA MANAGEMENT). In the first and larger group, reflecting the popularity of the out-of-the-box concept, randomness imposed on the thought process aiming at increasing the fluency and number of ideas; in the second group, the focus is on the quality and originality of ideas without considering their amount. The following review does not encompass all existing methods, but is intended to provide some flavor of contemporary creativity-enhancement methods.

Methods advocating unbounded randomness. The first broad group of techniques aims to identify a singular, unified thought process that leads individuals to generate successful creative ideas in different people. Its advocates collected written, first-person testimonies about the thought processes of creative people, and then attempts to define the general characteristics of the thought process that produces creative ideas. The results turned up countless inconsistencies between thought processes. Every creative person reported different experiences during his or her thinking, and no tangible, defining process that characterized creative thought was apparent. The researchers concluded that free thought and lack of coercion were the foundations of creative thinking – that breaking boundaries and laws, and using free association and intuitive tools, were the portals to creative thought.

Brainstorming. Most methods for the enhancement of idea generation devised over recent decades have been based on the belief that to find an original idea all we have to do is break away from existing (sound) mind frames and search for the surprising and the irregular, reaching the aspired goal of "generating a large number of ideas" (*see also* BRAINSTORMING). The implicit assumption behind such methods is that the greater the number of ideas, the greater the probability of achieving a set of quality ideas after filtering. Ideation is therefore measured often in quantitative rather than qualitative terms, and is directed in a random manner. Alex Osborn, founder of Batten, Barton, Durstine & Osborn (BBDO) was one of the first to suggest a creativity-enhancement technique – brainstorming. Brainstorming is based on a group of 6–10 people who focus on a problem. The cardinal rule is that judgment is deferred

and criticism prohibited: No line of inquiry is ruled out. The wilder the ideas that survive the better, for they might stimulate a new association that triggers more novel and useful ideas. Participants are encouraged to build upon ideas as they emerge, combine, and improve them. The atmosphere is positive; the objective of the brainstorming session is quantity, with the assumption that it leads to quality.

Following a first stage of divergence (conceptual brainstorming) – or what is termed by de Bono as "*messing around with ideas*," the method proceeds with an attempt at convergence (screening). During the second stage, the ideas (tens or sometimes hundreds) are filtered to produce a smaller set, which is subsequently examined and tested for economic feasibility and value, to produce a final set of novel, useful ideas.

Synectics, which means "bringing forth together" or "bringing different things into unified connection" was developed as a creative problem-solving method by Gordon. Synectics applies to the integration of diverse individuals into a purposeful group. This is an operational theory for the conscious use of the preconscious mechanisms present in human creative activity, and aims at increasing the probability of success in problem-stating, problem-solving situations (Gordon, 1961).

Lateral thinking. Lateral thinking is another widely used method popularized by De Bono (1971). The philosophy behind this approach may be conveyed by the analogy of digging a hole. While structural thinking is analogous to digging down in depth, lateral thinking is analogous to the search for a new spot to start digging. De Bono offers several provocative paths to force the thinker to consider different options for solution. Some of his techniques include inverting the situation, altering the situation to make it provocative, and considering interesting directions just because they are interesting (even where no benefit is seen).

Contrary to brainstorming, here the group does not take primacy; decision about the thinking path to be followed is more important. De Bono emphasizes that every creative idea is logical *a posteriori*, even if it looks illusive a priori. Many large organizations have reported successful cases of the use of lateral thinking, mostly for managerial problems.

Mind mapping. This method calls for a free association and flow of thoughts. An individual is instructed to draw a circle at the center of the paper and write in it a short description of the problem. He or she must then draw new circles around the page representing associations (*not* solutions) to the problem, which are linked to the first circle. Each new circle is now the origin for a new bundle of associations, which are noted again in new circles. When the paper is covered with circles (all linked to each other), the problem solver may look at the map and examine one small cluster of circles to try to find a solution. There are no reports of tests conducted to estimate the efficiency of this method.

Random stimulation. Sometimes presented as a complementary method to mind mapping, random stimulation posits that, if forced into the context of a problem, a remote analogy can sometimes stimulate a chain reaction of new thoughts and liberate a fixation. The method advocates choosing a random word (e.g., nail) and focusing thoughts on how that word could be part of a solution to the problem (e.g., devise a new schedule of production). The method is used mainly in writing and the arts. There are no reports of tests conducted to estimate the efficiency of this method.

Similar methods include *divergent thinking* (Guilford, 1970), in which thoughts flow in all directions from one starting point, even if the path seems illogical; *associative thinking* (Young, 1975), in which one puts together unrelated ideas; and *soft thinking* (Von Oech, 1989), which includes metaphorical, paradoxical, ambiguous, and fantasy thinking (giving the following direction, among others: "be foolish and silly").

All random-based methods are based on the requirement that judgment be suspended and divergent ideas emerge by associative thinking in an unconstrained space (Grossman, Rodgers, and Moore, 1988; Parnes, 1992). Therefore, the resultant idea-generation process is mostly random, or at least blind and haphazard (Campbell and Paller, 1989; Simonton, 1994).

How effective and efficient are randomness-based methods?. Methods based on *unbo-*

unded randomness are still often used in general management (Kiely, 1993; Rickards, 1998) as well as in the advertising industry (O'Guinn, Allen, and Semenik, 2000). Notwithstanding their popularity, they have been questioned in numerous studies (Bouchard, 1969; Connolly, Routhieaux, and Schneider, 1993; Diehl and Stroebe, 1987; Weisberg, 1992) Paulus *et al*. (1993), mainly because proponents of randomness-based methods have failed to provide research evidence of their effectiveness on the basis of a systematic assessment. In particular, they have failed to offer a method for controlling and discarding failed trials, thus dooming all trials and responses to qualitative commensurateness (Blachowicz, 1998, p. 21). One study (SOURCE) testing the performance of a group of problem solvers instructed to randomly "break the rules, get out of the square and shift paradigms," showed no significant differences between the ideas generated by this group and those generated by problem solvers given no instructions. Moreover, the study showed that while such methods may increase the apparent *novelty* of ideation, they decrease the *appropriateness*, *usefulness*, or *effectiveness* of the ideas produced.

As to the synergetic effect commonly identified with such unbounded randomness methods, which presupposes that a group of people thinking together is superior to a "nominal group" in which individuals think alone. It has been repeatedly and conclusively shown by investigators that the most prevalent method of *brainstorming* does not generate more ideas or greater creativity than do nominal groups (Diehl and Stoebe, 1987, 1991). At least one study asserted that this plays only a minor role in creativity ideation. In a controlled experiment, ideas suggested by individuals working alone were even evaluated as superior to those raised in brainstorming sessions (Weisberg, 1992). All in all, groups were shown to be suppressive of individual productivity, and the quality and originality of ideas generated by them to be inferior (Paulus, Ortega, and Brown, 1999; Sutton and Hargadon, 1996).

In the late 1980s, the theoretical foundation of the randomness methods also came under attack. Studies showed that an unrestricted scope does not necessarily lead to a treasure chest of creative ideas. In fact, restricted processes of thinking are more reliable for creativity (Bouchard, 1969; Connolly, Routhieaux, and Schneider, 1993; Diehl and Stroebe, 1987, 1991; Weisberg, 1992) (see below). In fact, it was further observed that people frequently asked to come up with new ideas sometimes try to find their own regulated means of becoming more productive at ideation tasks. They may, for instance, identify patterns common across different contexts and apply them on an ad hoc basis within a certain category. Such patterns will be less transient than the random extrication of thought (Boden, 1992; Dasgupta, 1994; Weisberg, 1992).

Methods advocating structure and regularity. Randomness-based methods of ideation enhancement are generally grounded in the belief that accidents and chance events enhance innovation. The belief that chance and irregular thinking enhance creativity is supported by many examples of new products created as a result of accidents or chance events. However, if we subscribe to this belief, we are ignoring the millions of unrecorded events in which accidents and chance events (or dreams) did not yield new products.

Underlying the methods advocating structure and regularity is a perception of creativity as an outcome of ordinary thinking that is only quantitatively different from everyday thinking, and does not necessarily require a qualitative leap (Perkins, 1981, 1988; Weisberg, 1992). Thus, according to Weisberg (1986), what distinguishes creative thinking from everyday thinking is its outcome rather than its manner of production. In support of this notion, Finke *et al*., found that constraining a problem by allowing fewer resources or components to be used increases the creativity of problem solvers (Smith, Ward, and Finke, 1992).

Compared to methods advocating randomness, methods advocating regularity – briefly reviewed below – manage cognitive processes rather than ideation sessions, are analytical and focused rather than random or blind, and are specific rather than general in applicability. Contrary to the "freedom" approach, the "structure" approach views creativity as a reproducible, learnable, and, in some cases, even potentially systematic thought process (Maimon

and Horowitz, 1999; Perkins, 1981, 1988; Simon, 1979, 1981). The thinking process is linear rather than random, and thus each step in the ideation process stems from the previous one.

TRIZ. In the 1940s, a chemical engineer named Genrich Altschuler postulated that underlying successful creative ideas and products are identifiable, repeated patterns or formulas. The existence of such patterns would obviate the need for searching the souls of inventors or tapping into the subjective thought process behind creative ideas (Altschuler, 1985, 1986). Altschuler's goal was to devise a systematic method to guide "ordinary" engineers toward creative solutions. By applying the principle of reverse engineering to the development process of more than 200 000 patents and technological inventions, he identified more than 40 patterns of invention, which he labeled *standards*. Standards were nonintuitive patterns that could be described, predicted, and controlled, independently of external influences: they consisted of system dynamics that could be determined solely by the intrinsic features of the products – a revolutionary idea in the field of creativity analysis.

Generations of Altschuler's students developed the method, and today it is widely used in many places. The focus of the method is still more technological although there are quite a few reported successes in many other fields. TRIZ methods and activity are reported in many web pages, although no studies on their application to marketing have been conducted.

Morphological analysis. The concept of *regularity* is embedded in a number of current ideation methods; for example, in marketing consider *morphological analysis*. This cluster of methods breaks down a system, outcome, or process into essential subconcepts, each representing one dimension in a multidimensional matrix. Ideas are created by searching the matrix for new, previously nonexistent combinations of attributes (Tauber, 1972). The major shortcoming of this cluster of methods is its lack of specific guidelines for combining the attributes, and its lack of a prescribed reduction mechanism facilitating the process of selecting the best ideas.

Creativity templates. Altschuler's principle has been applied to marketing, where recent creativity research has shown certain fundamental patterns or formulas underlying creative designs (Goldenberg, Mazursky, and Solomon, 1999a; Scott, 1994). In the context of innovation, Goldenberg, Mazursky, and Solomon (1999b) identified certain reoccurring regularities and found that these regularities represent the processes underlying new-product and innovation emergence. These regularities, termed *creativity templates*, were found to be identifiable, objectively verifiable, and generally observable. By virtue of their well-defined nature, templates may be described as step-by-step procedures that individuals can execute.

This approach is conceptually consistent with Altschuler's attempt to uncover the underlying patterns in the creative solutions of technological problems (Goldenberg and Mazursky, 2002). The main difference between Altschuler's approach and the creativity-template approach is the number and parsimony of patterns or templates proposed. Altschuler developed many. Researchers have found that most of the successful new product innovation development processes are covered by merely five distinct templates.

The creativity-template approach is a procedural ideation framework integrating two principles. First, there are several identified universal templates that underlie product evolution. These templates can serve to predict new candidate products by providing the context for ideation. Second, creativity is enhanced when thinking is channeled along predefined inventive routes. This principle is termed the *restricted scope principle*.

It was found (Goldenberg, Lehmann, and Mazursky, 2001; Goldenberg and Mazursky, 2002; Goldenberg, Mazursky, and Solomon, 1999a) that the application of templates may facilitate productive and focused ideation. Individuals who received training in creativity templates proposed more creative ideas than individuals who received training in total freedom methods, according to ratings by purpose-blind judges. Similar findings were obtained when the templates method was applied to creativity in advertising (Goldenberg, Mazursky, and Solomon, 1999b).

In this context of "constrained creativity," design structures *can* play the role of *attractors*: paths that the selforganized mind tends to follow, assisting the individual to process and organize information by using favorable processing routes proven in the past to lead to productive ideas (Kelso, 1997). The small number of paved routes (i.e., basic mental operations or mechanisms) avoids spending "a lot of time going down blind alleys" typical to brainstorming, and offers the much demanded "escape from freedom," which reduces anxiety (Fromm, 1941), thus maintaining – in Einstein's words – the "joy in creative expression and knowledge," and sustaining the "courage to create" (May, 1975).

Every year, hundreds of millions of dollars are invested in the development and launch of new products, most of which fail (Dahl, Chattopadhyay, and Gorn, 1999; Goldenberg, Lehmann, and Mazursky, 2001). It is argued that one of the several reasons for new-product failure is the absence of creativity (Sethi, Park, and Smith, 2001). Thus, understanding what makes a concept creative and how to cultivate creative ideation is especially crucial in the field of marketing. However, as Andrews and Smith (1996) argue, despite the importance of creativity in marketing programs, little is known about the factors that affect the generation of such programs; hence, marketing programs for many established products fall short in terms of creativity. Dahl and Moreau (2002) also note that empirical research examining the cognitive processes underlying the creation of novel product concepts is sparse.

Why is Creativity Important For Organizations?

Creativity and organizational innovation. The terms *creativity* and *innovation* are often used interchangeably. But even though the two constructs are conceptually related, it is important to distinguish between them. Creative performance refers to products and ideas produced at the individual level, whereas innovation refers to the successful implementation of new products at the organizational level (Amabile, (1988, 1983); Oldham and Cummings, 1996; Zhou, 2003). Innovation occurs when something new – something that has not been used or considered previously – is introduced to the organization. An innovation is a new artifact, product, system, technology, or process that has been successfully implemented in the organization.

Creative ideas or products do not always result in innovation. As Shalley, Zhou, and Oldham (2004) note, creativity refers to the development of novel and useful ideas, but only when these ideas are implemented successfully at the organization, can they be considered as innovations. Similarly, to introduce an innovation, the organization does not necessarily have to encourage or depend upon creativity – organizations can adapt products or processes that originated outside the organization (Woodman, Sawyer, and Griffin, 1993). Finally, sometimes an innovation is not surprising, nor original – it is just superior in some way. Innovations may represent an expected step in the evolution of a product, or may simply be moderately new.

Why is it important to study creativity to understand organizational innovation? As research often indicates, many innovations originate in creative ideas of individuals. Thus, although it is neither necessary nor sufficient, creativity is often the first step, the building block, of innovation (Amabile, (1988, 1997); George and Zhou, 2001; Tierney and Farmer, 1999; Zhou, 2003). Creativity is one of the major resources for innovation.

Organizational approaches to creativity enhancement. Past research on creativity often explored the impact of organizational-level constructs (e.g., organizational climate, Amabile *et al.*, 1995; perceived support, Scott and Bruce, 1995; presence of others, Shalley, 1995; goal setting, Shalley, 1991) or constructs reflecting stable individual differences (e.g., personality, Oldham and Cummings, 1996; problem-solving style, Scott and Bruce, 1995; intrinsic motivation, Amabile, 1997) (*see also* ORGANIZING FOR INNOVATION). However, more recently, researchers have argued that situational factors in organizations directly influence individuals' creativity, and this influence is moderated by various personal attributes (Amabile *et al.*, 1995; Amabile, 1988; Oldham and Cummings, 1996; Woodman, Sawyer, and Griffin, 1993). As a

result, two main approaches were considered to potentially lead to creativity in the organization. An individual perspective focuses on creativity as a personal attribute that may characterize a person to some degree. By selecting creative individuals as their members, organizations can increase creativity of their products and services. A situational perspective focuses on situational factors that may influence creativity: various organizational factors (e.g., reward systems, organizational values, and norms) may work independently or combined together to increase creativity (Lee *et al.*, 2004). Organizations can therefore increase individual creativity by shaping the environment and employing socialization and training processes that generate a creative-friendly environment.

More recently, however, researchers have pointed to the importance of studying effects of the interaction between personal and situational factors (George and Zhou, 2001; Oldham and Cummings, 1996; see review by Shalley, Zhou, and Oldham, 2004; Woodman, Sawyer, and Griffin, 1993; Zhou and George, 2003). Thus, for example, creativity was found to be associated with job complexity (Baer, Oldham, and Cummings, 2003; Oldham and Cummings, 1996) and job autonomy (see also Amabile *et al.*, 1995; Shalley, Gilson, and Blum, 2000 who focused exclusively on situational variables; Zhou and George, 2001).

Does creativity contribute to product success? Not only does creativity play a central role in the potential of firms to innovate, it also enhances their competitiveness in the global market (Miron, Erez, and Naveh, 2004) (*see also* SUCCESS FACTORS FOR NEW-PRODUCT DEVELOPMENT).

To be original, a product, by definition, has to be different from what exists. Thus increasing originality is likely to lead to feelings of surprise (Derbaix and Vanhamme, 2003), an emotion that is experienced when the perceived object does not map on to expectations (Reisenzein, 2000) – in this instance, the norms for the category. The question is, does originality affect market success? It has recently been argued that creative products that are both useful and original should have a high correlation with success and so they are expected to diffuse faster. However, so far there have not been any sufficiently convincing empirical results in favor of that argument. Contrary to what has been implied in the past (Carpenter, Glazer, and Nakamoto, 1994; Henard and Szymanski, 2001; Mishra, Kim, and Lee, 1996), originality alone may not be sufficient to ensure product success.

More directly, a meta-analysis by Henard and Szymanski (2001) showed that originality sometimes has a positive effect on performance and, at other times, the effect is negative. This led them to conclude that the effect of originality on performance is probably moderated by another unknown factor. Recent work by Derbaix and Vanhamme (2003) provides correlational evidence to suggest that feelings of surprise can lead to both positive and negative word of mouth (WOM). Thus, while increasing originality may lead to increasing levels of WOM, the valence of the WOM need not be positive; it can also be negative.

Goldenberg, Moldovan, and Chattopadhyay (2006) explored how new-product dimensions, specifically originality and usefulness, influence the WOM spread about the product and therefore its ultimate adoption and market size. They show that originality and usefulness have different roles in the adoption process of a new product, while originality amplifies both positive and negative WOM; usefulness determines the valence of WOM (positive or negative) and can increase market size. High originality, when combined with low usefulness, may produce high amounts of negative WOM and may lead to failure.

Another intervening factor that affects creativity's impact on product success may be timing. Being a first mover (Golder and Tellis, 1993) may not necessarily confer an advantage for a specific firm's strategy (Cooper, 1996; Lieberman and Montgomery, 1988, 1998). A follower advantage may be free riding, or even gaining a "gateway for a new entry" due to technology discontinuity. However the exclusive discovery of an emerging need is a commendable achievement and a necessary step toward a possible decision to take advantage of the opportunity to control market share through pioneering status.

Discussion

Many legends have been told about cases in which a person had developed an idea as a result of a dream. One famous story appears in almost every book dealing with creativity. Friedrich August von Kekule, who discovered and defined the formula of the paraffin molecule, later tried to decipher the secret of the benzene molecule. After many years of incessant research and thinking, something happened in his life: as he was sleeping by the hearth; he dreamt of balls jumping in the form of a snake swallowing its own tail. When he awoke he understood that this was the secret of benzene: contrary to what he had formerly believed, the benzene molecule is not "linear," but "circular" (hexagonal, to be exact). The lecture he delivered at the ceremony commemorating his discovery ended with the words: "Let us learn to dream, gentlemen." Koestler (1964) called this incident "probably the most important dream in history since Joseph's seven fat and seven lean cows."

By contrast, there are evidences that an excess of ideas and analogies obstructs the ideation process, and randomness and disorganization impede creativity. It has been realized that total freedom in problem solving is inadequate (Connolly, Routhieaux, and Schneider, 1993; Paulus *et al.*, 1993; Stroebe, Diehl, and Abakoumkin, 1991). "It follows," says Boden (1992), "that constraints – far from being opposed to creativity – make creativity possible. To throw away all constraints would be to destroy the capacity for creative thinking. Random processes alone, if they happen to produce anything interesting at all, can result only in first-time curiosities, not radical surprises."

An important part of being creative is to know the "rules of the game" and to become skilled at applying them (Perkins, 1981). Simonton (1984, 2003) states that the creative genius is an expert in a given domain, and is well acquainted with its rules and regularities. However, even in the exact sciences one cannot apply fixed rules mechanically. According to Aristotle, the matter of the *practical* – by being variable, particular, and perceptual in nature – is essentially indeterminate or indefinite. Practical wisdom must therefore use rules only as summaries and guidelines. It must itself be flexible, ready for surprise, prepared to see, and resourceful at improvisation or creativity (Kessels and Korthagen, 1996). MacIntyre (1977) aptly pointed to the interplay between *rule following* and *ruletranscending*: "Objective rationality is to be found in knowing how and when to put rules and principles to work and when not to. Because there is no set of rules specifying necessary and sufficient conditions for large areas of practices [such as creative advertising], the skills of practical reasoning are communicated only partly by precepts but much more by case-histories and precedents." The true work of the innovator consists in choosing among these combinations so as to eliminate the useless ones or rather to avoid the trouble of making them; and the rules that must guide this choice are extremely fine and delicate. It is almost impossible to state them precisely; they are (tacitly) felt rather than (explicitly) formulated.

Smith, Ward, and Finke (1992) note the existence of two distinct processing phases of creative thinking: a *generative* phase, followed by an *exploratory* one. These two phases are termed by Phillips (2005) as *inspiration* and *elaboration*. In Kelly's terms, there is a *creative cycle* of *loosening* and *tightening*. When being creative, we first *loosen* our constructions; then, finding a novel construction that seems to have some potential, we focus on it and *tighten* it, giving it substance or form (Kelly, 1955). An apparently similar rationale drives Barron (1969) to conclude that the essence of creativity is the ability to experience the extreme of psychological states – *crazy* and yet *sane;* in our view, creativity can be viewed as a constant oscillation between *surprise* and *regularity*.

Bibliography

Altschuler, G.S. (1985) *Creativity as an Exact Science*, Gordon and Breach Science Publishers.

Altschuler, G.S. (1986) *To Find an Idea: Introduction to the Theory of Solving Problems of Inventions*, Nauka, Novosibirsk.

Amabile, T.M. (1983) *The Social Psychology of Creativity*, Springer-Verlag, New York.

Amabile, T.M. (1988) A model of creativity and innovation in organizations. *Research in Organizational Behavior*, **10** (1), 123–167.

Amabile, T.M. (1997) Motivating creativity in organizations: on doing what you love and loving what you do. *California Management Review*, **40** (1), 39–58.

Amabile, T., Conti, R., Coon, H. *et al*. (1995) *Assessing the Work Environment for Creativity*, Division of Research Harvard Business School, Boston.

Anderson, H.H. (1959) *Creativity and its Cultivation, Addresses Presented at the Interdisciplinary Symposia on Creativity, Michigan State University, East Lansing*, Harper, New York.

Andrews, J. and Smith, D.C. (1996) In search of the marketing imagination: factors affecting the creativity of marketing programs for mature products. *Journal of Marketing Research*, **33** (2), 174.

Baer, M., Oldham, G.R., and Cummings, A. (2003) Rewarding creativity: when does it really matter? *The Leadership Quarterly*, **14** (4-5), 569–586.

Barron, F.X. (1969) *Creative Person and Creative Process*, Holt, Rinehart and Winston, New York.

Blachowicz, J. (1998) *Of Two Minds: The Nature of Inquiry*, State University of New York Press, Albany.

Boden, M.A. (1992) *The Creative Mind: Myths and Mechanisms*, BasicBooks, New York.

Boden, M.A. (1995) Creativity and unpredictability. *Stanford Humanities Review*, **4** (2), 123–139.

Bouchard, T.J. Jr. (1969) Personality, problem-solving procedure, and performance in small groups. *Journal of Applied Psychology*.

Bruner, J.R. (1992) AIDS and ERISA preemption: the double threat. *Duke Law Journal*, **41** (5) 1115–1156, http://www.jstor.org/stable/1372762.

Campbell, D.T. and Paller, B.T. (1989) Extending evolutionary epistemology to "Justifying" scientific beliefs (A sociological rapprochement with fallibilist perceptual foundationalism)," issues, *Evolutionary Epistemology*, State University of New York Press, Albany, pp. 231–257.

Carpenter, G.S., Glazer, R., and Nakamoto, K. (1994) Meaningful brands from meaningless differentiation: the dependence on irrelevant attributes. *Journal of Marketing Research*, **31** (3), 339–350.

Connolly, T., Routhieaux, R.L., and Schneider, S.K. (1993) On the effectiveness of group brainstorming: test of one underlying cognitive mechanism. *Small Group Research*, **24** (4), 490–503.

Cooper, R.G. (1996) New products, what separates winners from losers, *PDMA Handbook of New Product Development*, John Wiley & Sons, Inc., New York.

Dahl, D.W., Chattopadhyay, A., and Gorn, G.J. (1999) The use of visual mental imagery in new product design. *Journal of Marketing Research*, **36** (1), 18–28.

Dahl, D.W. and Moreau, P. (2002) The influence and value of analogical thinking during new product ideation. *Journal of Marketing Research*, **39** (1), 47–60.

Dasgupta, S. (1994) *Creativity in Invention and Design: Computational and Cognitive Explorations of Technological Originality*, Cambridge University Press.

De Bono, E. (1971) *Lateral Thinking for Management; A Handbook of Creativity*, American Management Association, New York.

Derbaix, C. and Vanhamme, J. (2003) Inducing word-of-mouth by eliciting surprise? A pilot investigation. *Journal of Economic Psychology*, **24** (1), 99–116.

Diehl, M. and Stroebe, W. (1987) Productivity loss in brainstorming groups: toward the solution of a riddle. *Journal of Personality and Social Psychology*, **53** (3), 497–509.

Diehl, M. and Stroebe, W. (1991) Productivity loss in idea-generating groups: tracking down the blocking effect. *Journal of Personality and Social Psychology*, **61** (3), 392–403.

Durgee, J.F., O'Connor, G.C., and Veryzer, R.W. (1996) *Using Mini-concepts to Identify Opportunities for Really New Product Functions*, Marketing Science Institute, Cambridge.

Field, R.C. (1968) Picnic by William Inge: a creative thesis in directing, Miami University, (Dept. of Communication and Theater).

Fromm, E. (1941) *Escape from Freedom*, Farrar & Rinehart, Inc., New York, etc.

George, J.M. and Zhou, J. (2001) When openness to experience and conscientiousness are related to creative behavior: an interactional approach. *Journal of Applied Psychology*, **86** (3), 513–524.

Gilford, J.P. (1967) *The Nature of Human Intelligence*, McGraw-Hill, New York.

Goldenberg, J. and Efroni, S. (2001) Using cellular automata modeling of the emergence of innovations. *Technological Forecasting and Social Change*, **68** (3), 293–308.

Goldenberg, J., Lehmann, D.R., and Mazursky, D. (2001) The idea itself and the circumstances of its emergence as predictors of new product success. *Management Science*, **47** (1), 69.

Goldenberg, J., Lowengart, O., Oreg, S., and Bar-Eli, M. (2010) How do revolutions emerge? Lessons from the Fosbury Flop, *International Studies of Management and Organization*, **40** (2).

Goldenberg, J. and Mazursky, D. (2002) *Creativity in Product Innovation*, Cambridge University Press.

Goldenberg, J., Mazursky, D., and Solomon, S. (1999a) Creativity templates: towards identifying the fundamental schemes of quality advertisements. *Marketing Science*, **18** (3), 333–351.

Goldenberg, J., Mazursky, D., and Solomon, S. (1999b) ESSAYS ON SCIENCE AND SOCIETY: creative sparks. *Science*, **285** (5433), 1495–1496.

Goldenberg, J., Moldovan, S., and Chattopadhyay, A. (2006) What Drives Word-of-mouth? The Roles of Product Originality and Usefulness, MSI working papers series, 06-111.

Goldenberg, J., Sagiv, L., Venkatraman, G., and Chattopadhyay, A. (2009) *Why Didn't I Think of That! Simplicity the Third Dimension of Product Creativity*, The Hebrew University of Jerusalem, Jerusalem.

Golder, P.N. and Tellis, G.J. (1993) Pioneer advantage: marketing logic or marketing legend? *Journal of Marketing Research*, **30** (2), 158–170.

Gordon, W.J.J. (1961) *Synectics, the Development of Creative Capacity*, Harper, New York.

Griffin, A. and Hauser, J.R. (1993) The voice of the customer. *Marketing Science*, **12** (1), 1–27.

Griffin, A. and Page, A.L. (1996) PDMA success measurement project: recommended measures for product development success and failure. *Journal of Product Innovation Management*, **13** (6), 478–496.

Grossman, S.R., Rodgers, B.E., and Moore, B.R. (1988) *Innovation, Inc.: Unlocking Creativity in the Workplace*, Wordware Publishing, Inc., Plano.

Gruber, H.E. and Wallace, D.B. (1999) The case study method and evolving systems approach for understanding unique creative people at work. *Handbook of Creativity*, Cambridge University Press, pp. 93–115.

Guilford, J.P. (1950) Creativity. *American Psychologist*, **5** (9), 444–454.

Guilford, J.P. (1959) Traits of creativity. *Creativity and its Cultivation*, Harper and Row, pp. 142–161.

Guilford, J.P. (1970) Creativity: retrospect and prospect. *Journal of Creative Behavior*, **7** (4), 247–252.

Guttman, L. (1968) A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika*, **33** (4), 469–506.

Haberland, G.S. and Dacin, P.A. (1992) The development of a measure to assess viewers? Judgments of the creativity of an advertisement: a preliminary study. *Advances in Consumer Research*, **19**, 817–825.

Henard, D.H. and Szymanski, D.M. (2001) Why some new products are more successful than others. *Journal of Marketing Research*, **38** (3), 362–375.

Jackson, P.W. and Messick, S. (1965) The person, the product, and the response: conceptual problems in the assessment of creativity. *Journal of Personality*, **33** (3), 309–329.

Kasof, J. (1995) Explaining creativity: the attributional perspective. *Creativity Research Journal*, **8** (4), 311–366.

Kelly, G. (1955) *The Psychology of Personal Constructs*, Norton, New York.

Kelso, J.A.S. (1997) *Dynamic Patterns: The Self-organization of Brain and Behavior*, MIT Press, Cambridge.

Kessels, J.P.A.M. and Korthagen, F.A.J. (1996) The relationship between theory and practice: back to the classics. *Educational Researcher*, **25** (3), 17–22.

Kiely, T. (1993) The idea makers. *Technology Review, Manchester, NH*, **96** (1), 32–40.

Koestler, A. (1964) *The Act of Creation*, Macmillan, New York.

Kosko, B. (1992) *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*, Prentice Hall, Englewood Cliffs.

Lee, F., Edmondson, A.C., Thomke, S., and Worline, M. (2004) The mixed effects of inconsistency on experimentation in organizations. *Organization Science*, **15** (3), 310–326.

Lieberman, M.B. and Montgomery, D.B. (1988) First-mover advantages. *Strategic Management Journal*, **9**, 41–58.

Lieberman, M.B. and Montgomery, D.B. (1998) First-mover (dis)advantages: retrospective and link with the resource-based view. *Strategic Management Journal*, **19** (12), 1111–1125.

Lubart, T.I. (1995) Creativity, in *The Nature of Creativity: Contemporary Psychological Perspectives* (ed. R.J. Sternberg), University Press, Cambridge.

Lumsden, C.J. (1999) Evolving creative minds: stories and mechanisms, in *Handbook of Creativity* (ed. R.J. Sternberg), Cambridge University Press, Cambridge, New York.

MacIntyre, A.C. (1977) Epistemological crises, dramatic narrative, and the philosophy of science. *The Monist*, **60** (4), 453–472.

Mackinnon, D.W. (1970) The nature and nurture of creative talent. *American Psychologist*, **17**, 484–495.

Maimon, O.Z. and Horowitz, R. (1999) Sufficient conditions for inventive solutions. *Systems, Man and Cybernetics, Part C, IEEE Transactions on*, **29** (3), 349–361.

Marschak, T.A., Glennan, T.K., and Summers, R. (1967) *Strategy for R & D; Studies in the Microeconomics of Development*, Springer-Verlag, New York.

Martindale, C. (1989) Personality, situation, and creativity, *Handbook of Creativity*, Cambridge University Press, pp. 211–232.

Maslow, A.H. (1959) Creativity in self-actualizing people, in *Creativity and Its Cultivation* (ed. H.H. Anderson), Harper, New York.

May, R. (1975) *The Courage to Create*, Norton, New York.

Mayer, R.E. (1999) *Fifty Years of Creativity Research*, Cambridge University Press, London.

Miron, E., Erez, M., and Naveh, E. (2004) Do personal characteristics and cultural values that promote innovation, quality, and efficiency compete or complement each other? *Journal of Organizational Behavior*, **25** (2), 175–199.

Mishra, S., Kim, D., and Lee, D.H. (1996) Factors affecting new product success: cross-country comparisons. *Journal of Product Innovation Management*, **13** (6), 530–550.

Mozart, W.A. (1952) in *"A Letter," in The Creative Process: A Symposium* (ed. B. Ghiselin), University of California Press, Cambridge University Press.

Mumford, M.D. and Gustafson, S.B. (1988) Creativity syndrome: integration, application, and innovation. *Psychological Bulletin*, **103** (1), 27–43.

O'Guinn, T.C., Allen, C.T., and Semenik, R.J. (2000) *Advertising*, South-Western College Publishing, Cincinnati.

Oldham, G.R. and Cummings, A. (1996) Employee creativity: personal and contextual factors at work. *The Academy of Management Journal*, **39** (3), 607–634.

Osche, R.A. (1990) *Before the Gates of Excellence: The Determinants of Creative Genius*, Cambridge University Press, Cambridge, New York.

Parnes, S.J. (1992) *Source Book for Creative Problem-solving: A Fifty Year Digest of Proven Innovation Processes*, Creative Education Foundation Press, Buffalo.

Paulus, P.B., Dzindolet, M.T., Poletes, G., and Camacho, L.M. (1993) Perception of performance in group brainstorming: the illusion of group productivity. *Personality and Social Psychology Bulletin*, **19** (1), 78–89.

Paulus, P.B., Ortega, A.H., and Brown, V. (1999) Group creativity, in *Social Creativity* (eds A. Montuori, and R.E. Purser), Hampton Press, Cresskill.

Perkins, D.N. (1981) *The Mind's Best Work*, Harvard University Press, Cambridge.

Perkins, D.N. (1988) The possibility of invention. *The Nature of Creativity*, Cambridge University Press,- pp. 362–385.

Phillips, H. (2005) Looking for inspiration. *New Scientist (1971)*, (2523), 40–42.

Poincaré, H. (1929) *The Foundations of Science: Science and Hypothesis, the Value of Science, Science and Method*, The Science Press, New York.

Poincaré, H. (1952) *Science and Hypothesis*, Dover.

Ray, M.L. and Myers, R. (1986) *Creativity in Business*, Doubleday, Garden City.

Reisenzein, R. (2000) The subjective experience of surprise, in *The Message Within: The Role of Subjective Experience in Social Cognition and Behavior* (eds H. Bless, and J.P. Forgas), Psychology Press, Philadelphia.

Rickards, T. (1998) Assessing organisational creativity: an innovative benchmarking approach. *International Journal of Innovation Management*, **2** (3), 367–382.

Rogers, C. (1959) Toward a theory of creativity, in *Creativity and Its Cultivation* (ed. H.H. Anderson), Harper & Row, New York.

Rothenberg, A. and Hausman, C.R. (1976), *The Creativity Question*, Duke University Press, Durham.

Runco, M.A. and Charles, R.E. (1993) Judgments of originality and appropriateness as predictors of creativity. *Personality and Individual Differences*, **15** (5), 537–546.

Scott, L.M. (1994) Images in advertising: the need for a theory of visual rhetoric. *Journal of Consumer Research*, **21** (2), 252.

Scott, S.G. and Bruce, R.A. (1995) Decision-making style: the development and assessment of a new measure. *Educational and Psychological Measurement*, **55** (5), 818.

Sethi, R., Park, C.W., and Smith, D.C. (2001) *Cross-functional Product Development Teams and the Innovativeness of New Consumer Products*, Cambridge University Press, Cambridge.

Shalley, C.E. (1991) Effects of productivity goals, creativity goals, and personal discretion on individual creativity. *Journal of Applied Psychology*, **76** (2), 179–185.

Shalley, C.E. (1995) Effects of coaction, expected evaluation, and goal setting on creativity and productivity. *Academy of Management Journal*, **38** (2), 483–503.

Shalley, C.E., Gilson, L.L., and Blum, T.C. (2000) Matching creativity requirements and the work environment: effects on satisfaction and intentions to leave. *The Academy of Management Journal*, **43** (2), 215–223, http://www.jstor.org/stable/1556378.

Shalley, C.E., Zhou, J., and Oldham, G.R. (2004) The effects of personal and contextual characteristics on creativity: where should we go from here? *Journal of Management*, **30** (6), 933–958.

Shye, S. and Elizur, D. (1994) *Introduction to Facet Theory: Content Design and Intrinsic Data Analysis in Behavioral Research*, Sage Publications, Thousand Oaks.

Simon, H.A. (1979) *Models of Thoughts*, Yale University Press.

Simon, H.A. (1981) The Sciences of the Artificial.

Simonton, D.K. (1984) *Genius, Creativity, and Leadership: Historiometric Inquiries*, Harvard University Press, Cambridge.

Simonton, D.K. (1994) Individual differences, developmental changes, and social context. *Behavioral and Brain Sciences*, **17**, 552–553.

Simonton, D.K. (2003) Scientific creativity as constrained stochastic behavior: the integration of product, person, and process perspectives. *Psychological Bulletin*, **129** (4), 475–494.

Sinnott, E.W. and Anderson, H.H. (1959) The creativeness of life, *Creativity and its Cultivation*, Harper & Row, New York.

Smith, G.F. (1998) Idea-generation techniques: a formulary of active ingredients. *The Journal of Creative Behavior*, **32** (2), 107–133.

Smith, S.M., Ward, T.B., and Finke, R.A. (1992) *Creative Cognition: Theory, Research, and Applications*, MIT Press.

Smith, R.E. and Yang, X. (2004) Toward a general theory of creativity in advertising: examining the role of divergence. *Marketing Theory*, **4** (1/2), 31–58.

Sobel, R.S. and Rothenberg, A. (1980) Artistic creation as stimulated by superimposed versus separated visual images. *Journal of Personality and Social Psychology*, **39** (5), 953–961.

Stein, M.I. (1953) Creativity and culture. *Journal of Psychology: Interdisciplinary and Applied*, **36**, 311–322.

Sternberg, R.J. (1985) *Beyond IQ: A Triarchic Theory of Human Intelligence*, Cambridge University Press, Cambridge Cambridgeshire, New York.

Sternberg, R.J. (1988) *The Nature of Creativity: Contemporary Psychological Perspectives*, Cambridge University Press, Cambridge, New York.

Sternberg, R.J. and Lubart, T.I. (1996) Investing in creativity. *The American Psychologist*, **51** (7), 677–688.

Stroebe, W., Diehl, M., and Abakoumkin, G. (1991) *The Illusion of Group Effectivity*, Psychologisches Institut, Tübingen.

Sutton, R.I. and Hargadon, A. (1996) Brainstorming groups in context: effectiveness in a product design firm. *Administrative Science Quarterly*, **41** (4), 685–718.

Tauber, E.M. (1972) HIT: heuristic ideation technique-a systematic procedure for new product search. *Journal of Marketing*, **36** (1), 58–61.

Taylor, C.W., Smith, W.R., and Ghiselin, B. (1963) The creative and other contributions of one sample of research scientists, in *Scientific Creativity: Its Recognition and Development*, John Wiley & Sons, Inc., New York.

Tellis, G.J. (1998) *Advertising and Sales Promotion Strategy*, Addison Wesley, Reading.

Thorson, E. and Zhao, X. (1997) Television viewing behavior as an indicator of commercial effectiveness, in *Measuring Advertising Effectiveness* (ed. W.D. Wells), L. Erlbaum Associates, Mahwah.

Tierney, P. and Farmer, S.M. (1999) An examination of leadership and employee creativity: the relevance of traits and relationships. *Personnel Psychology*, **52** (3), 591–620.

Ulrich, K.T. (1988) Computation and Pre-Parametric Design. Massachusetts Institute of Technology, Cambridge Artificial Intelligence Laboratory, Defense Technical Information Center, Ft. Belvoir.

Unsworth, K. (2001) Unpacking creativity. *The Academy of Management Review*, **26** (2), 289–297.

Urban, G.L. and von Hippel, E. (1988) Lead user analyses for the development of new industrial products. *Management Science*, **34** (5), 569–582.

Von Hippel, E. (1984) Novel product concepts from lead user: segmenting users by experience. *MSI Report*, **84**, 109.

Von Hippel, E. (1989) New product ideas from lead users. *Research Technology Management*, **32** (3), 24–27.

Von Oech, R. (1983) *A Whack on the Side of the Head: How to Unlock Your Mind for Innovation* (Warner Books ed.), Warner Books, New York.

Wallas, G. (1926) *The Art of Thought*, Harcourt, Brace and Company, New York.

Weisberg, R.W. (1986) *Creativity: Genius and Other Myths*, W.H. Freeman, New York.

Weisberg, R. (1992) *Creativity: Beyond the Myth of Genius*, W H Freeman, New York.

Wells, W. (1989) *Planning for R.O.I.: Effective Advertising Strategy*, Prentice Hall, Englewood Cliffs.

Woodman, R.W., Sawyer, J.E., and Griffin, R.W. (1993) Toward a theory of organizational creativity. *The Academy of Management Review*, **18** (2), 293–321.

Woodworth, R.S. (1938) *Experimental Psychology*, Henry Holt and Company, Inc., New York.

Young, J.W. (1975) *A Technique for Producing Ideas*, Crain Communications, Chicago.

Young, J.G. (1985) What is creativity. *The Journal of Creative Behavior*, **19** (2), 77–87.

Zhou, J. (2003) When the presence of creative coworkers is related to creativity: role of supervisor close monitoring, developmental feedback, and creative personality. *Journal of Applied Psychology*, **88** (3), 413–422.

Zhou, J. and George, J.M. (2001) When job dissatisfaction leads to creativity: encouraging the expression of voice. *The Academy of Management Journal*, **44** (4), 682–696.

Zhou, J. and George, J.M. (2003) Awakening employee creativity: the role of leader emotional intelligence. *The Leadership Quarterly*, **14** (4-5), 545–568.

product platforms

Marc H. Meyer, Olivier de Weck, and Tucker Marion

Basic Definitions

It has been shown that a key to success is to leverage existing know-how and assets across existing products as well as new market applications (Meyer, Tertzakian, and Utterback, 1997). Designing, developing, and integrating "product platforms" is a tried and true practice for accomplishing that end (Meyer and Lehnerd, 1997). We first define the meaning and context for product platforms in new-product development, and then provide a method for planning product platforms and linking them to current or new-product portfolios.

Definitions for product platforming are very important because "platform" is now used in so many different ways. To establish a meaning that we find most useful in the practice of new-product line development, let us begin with Figure 1. That figure provides an approach for leveraging product platforms to current and new market applications in the form of distinct product lines with specific products (*see* PRODUCT-LINE STRATEGIES).

Market segments and applications. At the top of the figure is a representation of a firm's current and potential target market segments, with a range of distinct *market applications*. A market segment is a specific group of users, made specific by virtue of classic segmentation criteria (geographic location, income level, gender, or industry affiliation) and behavior criteria (a high-performance user versus a price buyer, for example). Market applications are formed by applying potential product *uses* against these different users. A market application is therefore the intersection of specific groups of *users* and specific types of product *uses*.

In the figure, different market applications are labeled A, B, C, and D. Each represents distinctly different needs and user preferences and mainly requires different functionality in some shape or form in the products or services designed to address the target applications.

To illustrate, an automobile manufacturer can segment potential users in terms of age, gender, and family status – for example, as families, mature professionals, or young adults. These users then have different product uses: family driving, commuting to work, driving as part of a trade, or driving for social experience (the youth market). From the intersection of these users and uses come specific market applications: a family van, an upscale commuter car, a pickup truck, and/or a stylish sports utility vehicle (SUV) designed for Gen Y.

Product portfolio. The aggregation of all the products within the product lines intended to address a firm's target market applications can be referred to as the firm's *product portfolio* (Wheelwright and Clark, 1992; McGrath, 2001) (*see also* PORTFOLIO MANAGEMENT). In Figure 1, the product portfolio contains all the specific product variations within the product lines serving market applications A, B, C, and D. The product portfolio specifies different product offerings, typically at different price points, for different target applications – with an eye on beating competition, increasing demand, and achieving overall operating profit. These specific offerings can be arrayed on the same market segmentation grid. Rather than just "A" in the figure, a firm might offer several or more specific products for a specific market application. The firm might also not play in that market space because of competitive intensity or lack of market knowledge. Staging an entry into specific market applications over a period of years is another possibility.

Creating robust product portfolios requires deep insight into the needs, preferences, and aspirations of different potential types of customers (*see* VOICE OF THE CUSTOMER). The marketing goal of effective portfolio planning is that each element of the product portfolio brings complete satisfaction – an even better, delight to each specific target user and his/her intended product use along the price/performance/feature or capability continuum. Portfolio planning should also establish the clear streams of revenue derived from specific product offerings for target market applications. In other words, the product portfolio becomes a revenue map.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

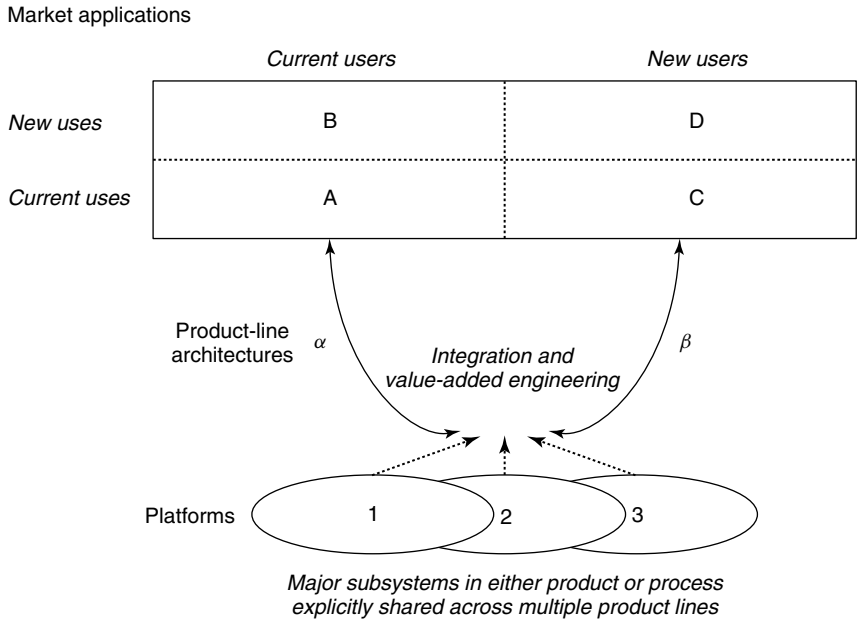


Figure 1 Leveraging platforms to new market applications.

Product-line architecture. Figure 1 also represents a firm that applies major common subsystems to its different market applications. An example might be a common engine across different passenger automobiles and sports utility vehicles (Honda), or a common database management system serving different software modules within enterprise bookkeeping and customer management (SAP). The ability to leverage these common subsystems relies on a clear architecture for each product within the product portfolio. That architecture defines not only all the subsystems but also the interfaces between these subsystems.

Robust product-line architecture allows an engineering team to combine subsystems common to the entire product line with subsystems that are unique to specific product variations in a manner that is both effective and efficient. It also allows a team to produce different levels of functionality within the product line without reengineering each individual offering – typically seen as "good, better, best" variations derived from the same product-line architecture (Meyer and Lehnerd, 1997). In addition, a product-line architecture may in fact span multiple market applications – A, B, C, and D taken as whole in Figure 1 – or each market application may require its own unique product-line architecture.

In Figure 1, we show two product-line architectures, labeled *alpha* and *beta*. These represent distinct configurations of subsystems and the interfaces between those subsystems required for the overall construction or assembly of products created from the product-line architecture. Continuing our example, Honda has distinct product architectures serving its Civic and Accord product lines, and another product-line architecture for its entry level SUVs, its large SUVs, and its minivans. A firm must periodically renew its product-line architectures to incorporate new technology breakthrough – such as using new hybrid engine technology in passenger cars.

Product platforms. *Product platforms* shown in the Figure 1 as 1, 2, and 3 represent specific modules of functionality. These are modules used within and across different product-line

architectures. A single subsystem shared across multiple products is considered a product platform, and across multiple product lines, truly powerful and robust. A product platform is an operational building block for a range of final products. For example, when a razor manufacturer develops a shaving blade of the same dimensions, coating, and edge for different products, it has created a powerful product platform. The same user interface across a suite of software products brings complete satisfaction to the user, simplifies training, and decreases switching costs (Cooper, 2004).

A product platform can also be an interface between different subsystems within a product-line or multiple product-line architectures. In software, .Net is a product platform enabling distributed computing in Microsoft's own operating systems, database, portal, and content management tools, its Office application suite, and third-party software products. Common product platforms – the core subsystems within an architecture and the interface between these and other nonshared subsystems – lead to efficient use of engineering resources, economies in materials purchase, and more rapid cycle time for developing new products that can use the common building blocks.

Process platforms. The manufacture, assembly, or fulfillment of a product platform is typically creating a shared asset that is used to produce different specific products within a product line (*see also* PROCESS INNOVATION). That shared asset may be considered a *process platform*. Common process platforms lead to substantial advantages in terms of capital for plant and equipment (where costs are shared across multiple product lines), as well as efficiency and learning in actual production. For example, Honda uses the same assembly line for manufacturing its Civics passenger lines and Element sport utility vehicle; and Mars, Inc. produces its heart-healthy CocoaVia chocolate bars on the same manufacturing asset as its Kudos snack bar (Meyer, 2007). Flexibility in the shared manufacturing asset to produce different configurations is an essential design goal (Hauser and de Weck, 2007).

Working with Product Platforms: Defining Fixed Points and Flex Points Within a Product-Line Architecture

Figure 2 shows how a team can think about the combination of product platforms with other subsystems within a product-line architecture. In the figure, each product line has its own distinct product-line architecture (defined as the major subsystems and interfaces that are the foundations of all the specific products within that product line of business). Over time, that architecture evolves to meet new user requirements and embrace new core technology. In the figure, that architecture is depicted as layers of subsystem technology – some common, others unique, to specific product lines within the product portfolio.

In platform-focused firms, the value-added engineering performed for individual products is a portion of the total engineering effort, because the most of the technology for single products comes from common existing platforms. The power of the platform discipline becomes evident when a substantial percentage of all the modules within and between product-line architectures are common. There can be cases where the degree of commonality is 100%, and variation for different groups of customers is achieved through *services* provided with the product. However, it has also been shown that taking platforming too high up the stack can make differently branded products look and function all too similarly – saving on cost of goods but disappointing users (Simpson, 2004). The more typical scenario, however, is that half or more of the engineering required for a new-product development within an established product line is achieved by the integration of preexisting subsystems based on preexisting interfaces.

With time, the functionality, performance, and cost of common subsystems are improved by the firm itself or its suppliers. Robust interface design allows engineers to migrate to increasingly powerful generations of subsystem technology without violating the overall product-line architecture. The ability to replace aging components with newer, better ones comes from effective technology roadmapping (Albright and Kappel, 2002).

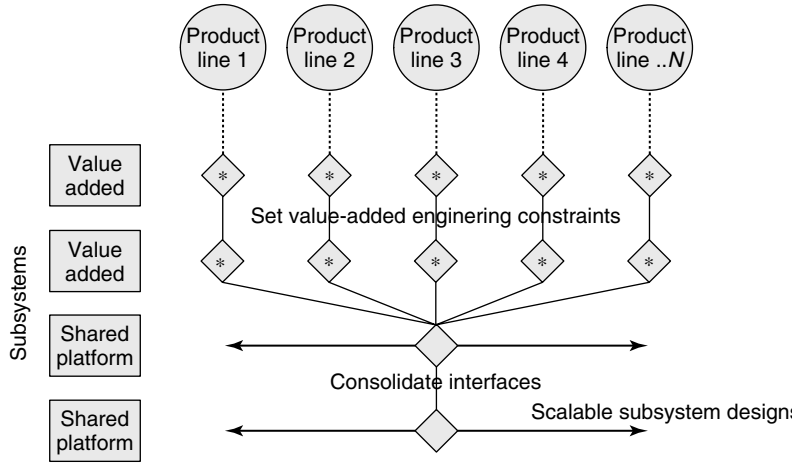


Figure 2 Fix points and flex points in design: combining common product platforms with unique subsystems.

Interface discipline is essential for successful product platforming (Boothroyd and Dewhurst, 1987; Sundgren, 1995; Cooper, 2004; Meyer and Webb, 2005). This concept is represented by the words *single channel interfaces to and from subsystems* in Figure 2. This means one clear method or process for connecting to that subsystem from any other subsystem, clearly documented and strongly enforced. Unfortunately, as engineers add new functionality to other subsystems, they also tend to create new custom interfaces between these subsystems. Over time, a multiplicity of interfaces becomes a tremendous impediment to the innovation of a new subsystem and a value-added product development. Without interface discipline, it becomes increasingly expensive to make even the most incremental of product improvements. On the other hand, robust interface design can enable greater levels of product variation at minimal cost. For example, earthmoving equipment manufacturers, such as Caterpillar, have designed flexible coupling devices to rapidly attach different work tools (buckets, hammers, and forklifts) to the same machine – such as a backhoe loader.

Planning for Scalability in Product and Process Platforms

The primary design requirement for a product or process platform must be the flexibility to fit within different product-line architectures, and scalability, to support increasingly powerful products. This enables a firm to efficiently leverage these common modules – its product platforms and the production processes for them – to new market applications. The process of mapping specific product platforms, with varying levels of functionality, to specific products is well described by Thomson and Broms (2000) with the example of how Scania designs its trucks.

Defining the range of functionality for a common subsystem or interface, for example, the product platform, requires that a team first have a working definition of the product portfolio. That portfolio may include all the products planned with a single product line, or all the products planned for multiple product lines depending on the scope and focus of the development initiative (as represented earlier in Figure 1). With the portfolio in hand, the team must then carefully study user requirements in terms of performance, other types of functionality, and cost – typically arrayed as levels of performance/price within the product line(s) serving the target applications. These functionality requirements (*see* PRODUCT SPECIFICATIONS) may be strikingly different for different markets – male versus female; developed country versus emerging market; enterprise software versus home-use software

market. With requirements in hand, the team can then design its product platforms to serve its target applications. By achieving this, the team creates a *scalable* product platform (de Weck, 2006).

Figure 3 shows one classic instantiation of engineering a scalable product platform – in this case, a common motor that could deliver a broad range of power based on varying wire stack length, and which served levels of "good, better, best" for Black and Decker's drills, sanders, jig saws, circular saws, and other major power tools product lines. Figure 4 shows a more recent example, Honda's family of engines that have served "good, better, best" versions of both its passenger cars and sports utility vehicles. Scalable product platforms such as these, and their associated high-volume manufacturing processes, are powerful enablers to reduce cost of goods on a per unit basis and to meet rapidly expanding levels of demand for products in the line.

Product Platforms Have Implications beyond Product and Technology

Making the business case for product platforms is increasingly important in resource-constrained business environments where the benefits of a platform program must be weighed against more traditional single-product developments. Experience has shown that in many industries, it is initially more expensive to design and develop a robust, scalable product platform and its associated manufacturing processes. To be able to justify the benefit of the platform effort to the business, a team must develop a business case.

The business plan for platforming should aggregate operating profits projected for all products in the platform-based portfolio and use these aggregated sums in net present value and/or ROI calculations. A team might find, for example, that the investment required to develop platforms might be more than covered by the lower cost of goods from such commonality. However,

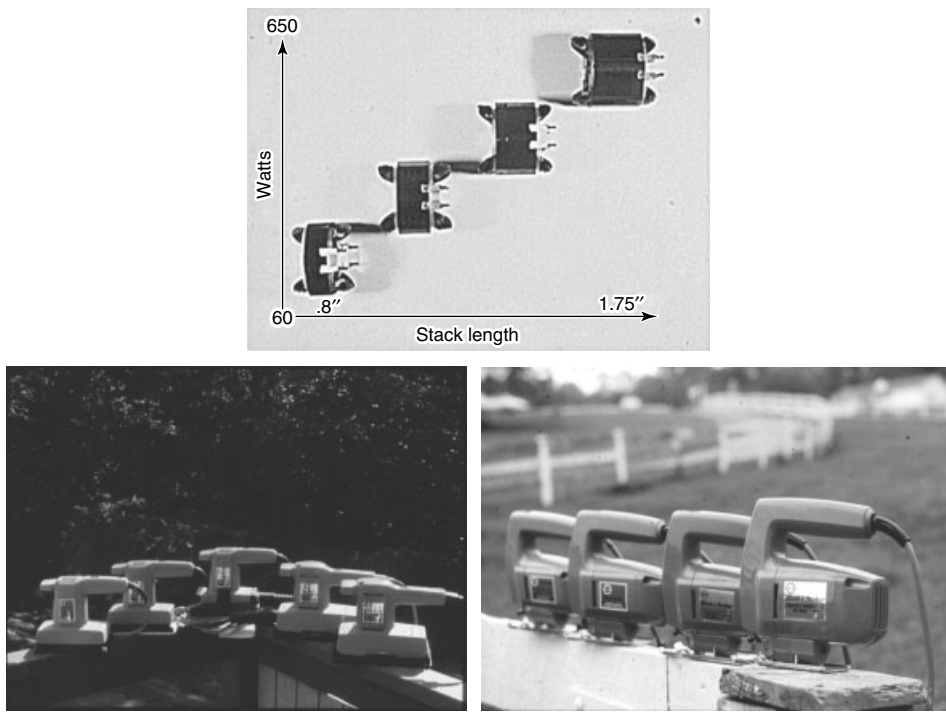


Figure 3 Black and Decker achieves magic with its scalable motors – the product platform shown incorporated in two of many product lines (reprinted from Meyer and Lehnerd, *The Power of Product Platforms*, 2007).

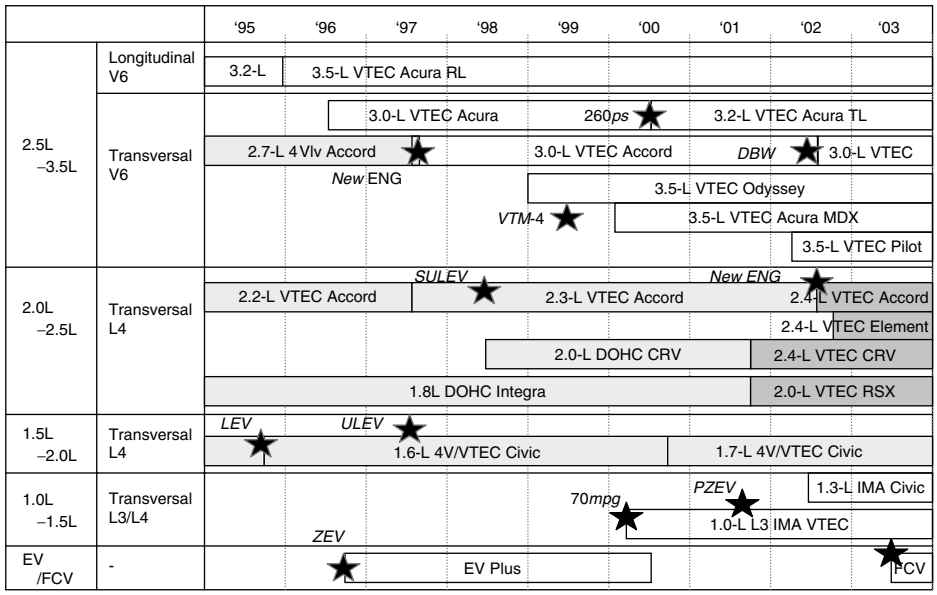


Figure 4 An engine for growth in Honda: a road map of common power trains from 1995 to 2003 (reproduced with permission from Honda Motor Company, Ltd., reprinted from Meyer, The Fast Path to Corporate Growth 2007).

arriving at this type of detailed understanding of material and conversion costs for platform-based product lines requires a dedicated effort to do well. Clearly, a platform-planning team should include individuals with strong financial acumen who know how to gather cost information and integrate these with pro forma income statements and capital plans.

Breaking cost barriers typically requires a team to challenge the conventional thinking within the company in terms of manufacturing processes and suppliers. If a team is bold in its thinking, it can deliver more powerful products at lesser cost than the firm's current offerings, and its proposal to senior management becomes much more attractive. For example, it can recommend lower market prices for its new products, increasing market share or in fact enlarging the current market to generate substantially greater revenue. Or, a team might recommend maintaining current market share and more slowly growing revenues, but substantially improved gross margin. Regardless of the final business-case approach, we encourage all platform teams not to simply seek product parity at lower cost, but rather, *to create better products that can also be produced at lower cost*. This helps achieve *value-cost leadership*, a foundation for sustained success (Meyer and Lehnerd, 1997).

Product platforms can also help a firm reconsider its business model. First, traditional premium manufacturers can apply platforming to lower their manufacturing costs, providing new mid-range or perhaps even low-end product offerings to match competition from offshore companies with lower labor costs. Many companies are at present considering platforming as a response to low-cost competitors who are scaling up their products to compete in mainstream markets. Second, platforming – particularly in software products – allows firms to create new streams of recurring revenue enabling plug-in modules for the installed base. Third, platforms can have beneficial effects for a firm's service business. Rather than training service staff on the inner workings of the many different subsystems and interfaces across an entire portfolio, a product line with shared subsystems allows economies in training and service delivery. Lastly, the design of services can themselves be platformed (*see* SERVICE INNOVATION MANAGEMENT),

focusing on common processes and measurements of quality and cost for those processes (Meyer and DeTore, 2001).

A business case must also consider that well-designed product and process platforms enable future innovation by providing a stable base upon which technology could be applied to create new market applications. Modifying the "flex points" and leveraging the "fixed points" within a product-line architecture is the key to creating niche-focused products at little cost and readying them for market testing. While one cannot fully anticipate the range of product variations in future years, the flexibility to respond to market opportunities should interest any forward thinking senior management team. Creating rich variety from simple means is a grounding principle of platform strategy.

There are, of course, many inspirational examples of successful platform management leading to business success. These include Honda, the different vehicles of which share power trains, and Boeing, whose 787 Dreamliner leverages scalable frame, cockpit, and interior subsystem technology across a broad range of formerly separate aircraft. Apple is another fine example. The iPOD, iTouch, and iPhone products share common operating systems and applications software platforms – including iTunes. Apple also broke tradition by making its iTunes software operational on Windows-based computers to broaden market reach. Aspiring to be like Honda, or Boeing, or Apple, however, is of directional value only; it is up to each company to determine how these principles apply best to its own specific situation. There are no short cuts.

To sum up, product platforms have intriguing implications that go well beyond products and technology. Readers must consider these to achieve business as well as technical success. Books such as *The Fast Path to Corporate Growth* (2007) and *Product Platform and Product Family Design: Methods and Applications* (2006) provide detailed methods for platforming new product lines and many examples of success.

Bibliography

Albright, R. and Kappel, T. (2002) Roadmapping in the corporation. *Research-Technology Management*, **42** (2), 31–40.

Boothroyd, G. and Dewhurst, P. (1987) *Product Design for Assembly*, Boothroyd & Dewherst, Inc., Wakefield.

Cooper, A. (2004) *The Inmates are Running the Asylum*, Sams Publishing, Indianapolis.

Hauser, D. and de Weck, O.L. (2007) Flexible parts manufacturing systems: framework and case study. *Journal of Intelligent Manufacturing*. **18** (3), 421–432.

Johnson, H.T. and Broms, A. (2000) *Profit Beyond Measure*, The Free Press, New York.

McGrath, M. (2001) *Product Strategy for High Technology Companies*, McGraw-Hill, New York.

Meyer, M.H. (2007) *The Fast Path to Corporate Growth: Leveraging Platforms to New Market Applications*, Oxford University Press, New York.

Meyer, M.H. and DeTore, A., (2001) Creating Platform-based Approaches to New services Development, *Journal of Product Innovation Management*, **18**, 188–204.

Meyer, M. and Lehnerd, L. (1997) *The Power of Product Platforms*, The Free Press, New York.

Meyer, M.H., Tertzakian, P., and Utterback, J. (1997) Metrics for managing product development within a product family context. *Management Science*, **43** (1), 88–111.

Meyer, M.H. and Webb, P. (2005) Modular, layered architecture: the necessary foundation for effective mass customization in software. *International Journal of Mass Customization*, **1** (1), 14–36.

Simpson, T.W. (2004) Product platform design and customization: status and promise. *Artificial Intelligence for Engineering Design*, **18**, 3–20.

Sundgren, N. (1995) Introducing interface management in new product family development. *Journal of-Product Innovation Management*, **16**, 40–51.

de Weck, O. (2006) Determining product platform extent, in *Product Platform and Product Family Design: Methods and Applications* (eds T.W. Simpson, Z. Siddique, and J. Jiao), Springer, New York, pp. 241–301.

Wheelwright, S. and Clark, K. (1992) *Revolutionizing New Product Development*, The Free Press, New York.

global product development

K. Sivakumar

Introduction, Definitions, and Organization

For our purposes, the word "global" means that the business activities of a firm span more than one country and, typically, several countries. We use the words *global*, *international*, and *cross-cultural* to denote similar concepts. Although the burgeoning growth of communication technology has resulted in what Friedman (2005) calls a *flat world*, in which location does not matter, globalization does have a spatial connotation. From a broader perspective, globalization encompasses variables such as language, culture, and time-zone commonality, as well as security issues. Globalization can manifest itself in several different ways: sales, sourcing, manufacturing, assembly, processing, accounting, and similar activities.

"Product development" refers to the process of conceiving of and creating a new product (Crawford and Di Benedetto, 2006); all new products must be "conceived, tangibly developed, and finally produced and sold in the market place" (Nakata and Sivakumar, 1996, p. 62). Product development involves several stages. The often interrelated stages can range from a two-part conceptualization of initiation and implementation (Zaltman, Duncan, and Holbek, 1973) to more than 10 stages, depending on how finely scholars want to categorize the different elements involved in product development (*see* THE STAGE-GATE IDEA TO LAUNCH SYSTEM). The outcome of product development is either a radical innovation or an incremental innovation. One way to conceptualize the innovation is a series of S-curves (Foster, 1986). Rising S-curves represent reduced costs and/or increased benefits along a given S-curve. Moving across the S-curve is radical innovation and moving along a given S-curve is incremental innovation (*see* TECHNOLOGY S-CURVE).

We define global product development as "a single, coordinated product development operation that includes distributed teams in more than one country utilizing a fully digital and connected collaborative product development process" (Eppinger and Chitkara, 2006, p. 23). "Product" is an omnibus term that includes goods (e.g., digital cameras) and services (e.g., customer service at call centers), including knowledge-based services (e.g., software development, pharmaceutical drug discovery) (*see* INNOVATION TYPOLOGIES).

The article is organized as follows. First, the antecedents and drivers of the increasing trend toward global product development are discussed; second, some major themes and domains of knowledge on the topic of global product development are highlighted; third, research on strategies for achieving success in global product development are delineated; finally, implications for research and managerial practice, including emerging opportunities and challenges in the domain of global product development, are discussed.

By definition, global product development is a multidisciplinary activity (*see* CROSS-FUNCTIONAL TEAM). Therefore, existing knowledge on the topic comes from various fields, including marketing, management, economics, research and development (R&D) management, international business, and management information systems. As Kleinschmidt, de Brentani, and Salomo (2007) note, the product development and globalization bodies of literature have been developing over the past several decades as extensive but separate streams, and there is a need to integrate the research in the two domains given the rapid increase in global product development. Given the space constraints here, the goal is to highlight some salient, representative, and emerging themes in this domain rather than to attempt to provide comprehensive coverage.

Drivers of Global Product Development

Globalization is an important trend of the modern world that will define our activities for several decades into the future. There are several manifestations of enhanced globalization, including the world's becoming "flat" (Friedman, 2005), the globalization of companies and customers, permanent or temporary movements or migrations of people

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

across countries, and the establishment of global monitoring organizations such as the World Trade Organization. There have always been debates among social scientists, policy makers, politicians, and academics about the appropriate level of globalization, the positive and negative effects of globalization, and related issues. Without taking a political position on the desirability or extent of globalization, it is clear that globalization is inevitable and that individuals and organizations are better served when they understand the antecedents, processes, and consequences of this trend, which enables them to exist and thrive in an increasingly complex and turbulent world. Global product development is one such inevitable manifestation in the operations of companies around the world. As Kleinschmidt, de Brentani, and Salomo (2007) argue, firms must commit to an integrated strategy of globalization and new-product development to be competitive in a business environment characterized by turbulence.

There are several trends that can be considered responsible for the increasing growth of global product development. First, the rapid globalization of companies, markets, and the employment base has led naturally to the globalization of product-development activity. Second, an increasingly educated workforce in developing countries such as India and China makes it convenient for companies to tap a wider pool of talent for all company operations, including product development. Third, intense competition in the marketplace and the need to reduce product development times provide an impetus for companies to use all available global resources to meet global customer needs. Fourth, the growth in information technology facilitates the knowledge development and sharing involved in new-product development across national borders (*see* LOCALE OF INNOVATION).

From media coverage and anecdotal evidence, it is clear that global product development is a phenomenon that pervades a wide range of industries, including computers (e.g., IBM), software (e.g., Microsoft), manufacturing (e.g., General Electric), consulting (e.g., Accenture), and pharmaceuticals (e.g., Merck). Although large multinational companies engage in global product development as a natural manifestation of their globalization, small and medium-size companies are involved in the process as well: as partners in strategic alliances, as upstream suppliers, or as downstream customers.

Also, countries such as India are leveraging human intelligence and knowledge rather than low-cost labor, which has been the traditional way to participate in international business. In addition, there is a qualitative change in the way globalized activities spur global product development; for example, in India plants that simply produce products according to specifications have a lesser role to play in product development than do dedicated design centers and R&D labs (*see* RESEARCH & DEVELOPMENT).

The Nature and Scope of Global Product Development

The scope of global product development varies depending on the nature of the product development task, including the time span, the degree of newness of the product, the number of people involved, and the importance of the product to the entities involved. At the low end, this might mean a small team of members from two units of a company in two different countries working together for a short-term, specific project to make a minor improvement to an existing product. At a more involved level, this might mean a large team from a network of companies working with their respective customers and suppliers in different tiers using elaborate communication and control systems to develop a radically new product that spans several years. Clearly, most global product development activities fall between the two extremes (e.g., a firm conducting its global new-product development with several subsidiaries via offshoring; a firm collaborating with an arm's-length foreign supplier and a local support office in a foreign country; a firm working to perfect a drug compound with an R&D unit in a foreign country; a firm's foreign subsidiary making a strategic new-product development alliance with another company to use its engineering or design capacity).

Therefore, the integration of company cultures and the integration of national cultures become important. For example, within-comp-

any global new-product development in a company such as IBM with operations in several countries represents fewer company culture differences and more national culture differences; global new-product development undertaken by a consortium of companies in two countries (e.g., the United States and the United Kingdom) represents more company culture differences and fewer national culture differences. Other nuances can be identified according to the nature of cultural difference (e.g., the cultural difference in a Canada–United States activity would be less than the cultural difference in a United States–India activity). The number of control systems that need to be established also varies on the basis of the type of product being developed and the nature of the relationship between the companies.

Much of the existing academic research has examined the process of product development (Crawford and Di Benedetto, 2006), and some of the core processes and routines carry over to global product development. Subramaniam, Rosenthal, and Hatten (1998) examined the processes and routines that firms use in global product development. They find that the processes that global companies follow (e.g., team composition) vary in the extent of the involvement of overseas subsidiaries. They also found industry-specific differences in how firms manage global product development. In particular, the authors found that with tacit knowledge about new-product design elements, firms employ cross-national teams and make use of overseas subsidiaries in the concept-development stage.

Business relationship factors such as buyer–seller interactions have significant influence on new-product development (Roy, Sivakumar, and Wilkinson, 2004). The buyer–seller factors become much more complicated in the context of global product development. Some of them may be simply a matter of scale (e.g., a larger community of practice), while others may represent qualitative differences (e.g., the extent and duration of communication moderated by the cultural dimensions).

The Role of National Culture in Global Product Development Teams

Almost all new-product development activity is carried out primarily by teams. Global new-product development teams were once considered the "next wave of corporate development" (Solomon, 1995, p. 50). "Like their domestic-only counterparts, global new-product development teams are typically cross-functional, commissioned to design and launch new products, and face significant budget, time, and other resource constraints. However, global product development teams are different in that members usually come from a range of countries and cultures. Values, orientations, and assumptions may radically differ among members" (Sivakumar and Nakata, 2003, p. 397). McDonough, Kahn, and Barczak (2001) found that more than half of the companies that they surveyed used globally dispersed teams in their product-development effort, a trend that is expected to continue. Global product development teams typically face greater challenges than do colocated teams (McDonough, Kahn, and Barczak, 2001). An important defining variable in this context is *national culture*, which is defined as the "collective programming of the mind which distinguishes the members of one group or category of people from those of another" (Hofstede, 1994, p. 4).

There are different conceptualizations of culture. Among the more popular and empirically based are those based on the work of Hofstede (1980) and his coresearchers. Nakata and Sivakumar (1996) examined the role of individualism, power distance, masculinity, uncertainty avoidance, and Confucian dynamic on the initiation and implementation stages of new-product development. They conclude that different national cultural dimensions have different effects on new-product development. For example, individualism facilitates the initiation stage of new-product development but inhibits the implementation stage. The effect of power distance, masculinity, and uncertainty avoidance are the reverse (higher levels of these factors impede the initiation stage but facilitate the implementation stage). Because of cultural differences, attention must be paid

to the design of global product development teams (Sivakumar and Nakata, 2003). The composition and management of a team must take into account factors such as the differential effect of cultural dimensions on the different stages of the new-product development process and the relative importance of the different stages in a given product-development task (Sivakumar and Nakata, 2003). Song and Parry (1997) have found that, although multicultural teams lead to more creativity and problem solving, communication breakdowns can occur, and the team leadership's responsibility in this domain becomes important.

The national culture dimensions also can influence other aspects of global product development such as R&D-marketing integration mechanisms. For example, Garrett, Buisson, and Yap (2006) argue that with high uncertainty avoidance, firms will rely more on formal planning and control tools to avoid uncertain situations.

Knowledge Management in Global Product Development

Clark and Fujimoto (1991) state that product development is an activity to reduce uncertainty because it generates and uses new knowledge; therefore, new-product development can also be visualized as information processing of organizations.

Clearly, product development is a knowledge-intensive process that involves constant accumulation, dissemination, and use of existing knowledge and the creation of new knowledge. New products involve new knowledge or information assets (Clark and Fujimoto, 1990). Whether the product-development task results in incremental innovation or radical innovation based on the new knowledge created (Tushman and Anderson, 1986), an important challenge in global product development involving different firms is knowledge management in general (Söderquist, 2006) and the management of intellectual property in particular (Roy and Sivakumar, in press) (*see* INTELLECTUAL PROPERTY RIGHTS). In addition, leveraging learning in the product-development process is complex, and attention must be paid to organization development and training (Duarte and Snyder, 1997). Knowledge management functions in global product development must recognize that different methods, such as a central knowledge management function, a project-decentralized knowledge management task force, and functionally located knowledge management cells all have different advantages and disadvantages, and that the optimal method depends on various internal and external factors related to the organization and the environment (Söderquist, 2006). Although the mix between tacit knowledge and explicit or codifiable knowledge is different across product-development tasks, the mix always involves challenges of knowledge management.

Given the complexity of structures, legal processes, and government regulations associated with global product development, the challenges of intellectual property management are unique. As Roy and Sivakumar (in press) point out, appropriate management of access, exploitation, and defense of intellectual property is governed by the extent of trust in the process and the amount of formal and informal control systems in place. By effectively managing intellectual property, firms will be able to achieve success in the generation of radically and incrementally new products (Roy and Sivakumar, in press).

Performance and Evaluation Issues in Global Product Development

The implicit objective of each firm is to optimize the performance of global product development to achieve long- and short-term success (*see* SUCCESS FACTORS FOR NEW-PRODUCT DEVELOPMENT). Crawford and Di Benedetto (2006) offer unique guidelines for evaluating a new-product development program. Such evaluations can be used for global product development and modified to account for the nature and scope of global product development. In general, the conventional measures of effectiveness and efficiency (e.g., ratio of new-product sales to total sales, product-development cycle time) can also help assess the performance of global product development, and many of the factors that govern the performance of product development in general influence the outcome of global product development.

Given the increased scope and complexity of the global product development process, some new challenges may arise. For example, when most of a new product's sales growth occurs in one or two countries, how should the firm understand the relative role of new-product sales of the global corporation? How can the differences in people's use of attitude scales across countries be reconciled with differences in the evaluation of people's assessment of new-product development programs?

There are some guidelines available from existing research on factors that influence global product development performance. The information-processing perspective of global product development implies that the fit between uncertainty and information-processing ability directly influences performance (Subramaniam, Rosenthal, and Hatten, 1998). Kleinschmidt, de Brentani, and Salomo (2007) found that performance in global product development is determined by organizational new-product development resources and new-product development process capabilities. Several researchers also offer operational guidelines for managing global product development. For example, Barczak and McDonough (2003) focused on effective ways to manage product-development teams and found that initial face-to-face meetings, increased communication among team members, and periodic progress meetings facilitate the success of global product development efforts. Wren, Souder, and Berkowitz (2000) studied the link between market orientation and product development in global firms. They found that both market intelligence gathering and customer orientation are related to superior new-product development performance. In addition, different cultural dimensions can have different influences in the performance of different stages of global product development (Nakata and Sivakumar, 1996).

Managerial and Research Implications

Although the phrase "necessity is the mother of invention" is often used as a cliché, the current economic, political, and security situation around the globe is likely to result in innovations in global new-product development procedures and processes. The following are some examples.

The differences in complexity between intra- and interfirm relationships involved in global product development will overlap considerably and become governed by interactions. For example, a scenario of constant interaction between two US companies may involve less complexity than the interaction between the US headquarters of a company and a branch unit in India or China, especially if there is frequent change of personnel in the branch office.

Second, although both codified knowledge and tacit knowledge are required and processed during global product development, improved information technology and communication devices will trend toward even more codified and routinized tacit knowledge. Relatedly, management of intellectual property will offer its own challenges.

Third, new advances in improving global product development are likely to come about in the case of knowledge-based services. Given the cultural differences among employees and customers, and the evolving nature of the knowledge economy, it is surprising that the political discourse in the United States (particularly during election time) is still dominated by the manufacturing economy, although we have long since moved from a predominantly manufacturing economy to a service economy and now even toward a knowledge economy. It will be a challenge for professionals to influence the public discourse to be consistent with the future-oriented evolution of the economy and the attendant new-product development process.

Similarly, the sensitive nature of globalization (represented by the phrase *outsourced jobs*) must change from an attitude of "us" versus "them" to mutual coexistence based on competitive advantages, local conditions, and process capabilities. Breaking down the barriers will invariably equalize the playing field, bring in more nuanced specializations as marketable resources, and bring about creative methods of collaboration.

Finally, for global product development, the company must recognize the global customer and also realize that that there are no more isolated markets in the world. Ultimately, global product

development is the domain that can effectively link people with knowledge, in addition to the traditional forms of monetary exchange or manufacturing facilities, as a resource.

Conclusion

Global product development represents the intersection of the two most important business activity domains in the foreseeable future. Although the nature of global product development can be considered an extension of product development in general and some of its aspects may be similar to traditional product development (e.g., the types of activities undertaken, the need for multidisciplinary teams), the complexities and nuances of global product development can be very different (e.g., modes of communication, number of people involved, location of different teams, cultural differences). Firms must proactively address such complexities and nuances. Increased global product development will result in the evolution of the locus of the product-development activity, the types of products being developed, and areas in which the products are registered and marketed. Thus, there will be a need to focus on the global dimension in all product-development research and practice.

Bibliography

Barczak, G. and McDonough, E.F. III (2003) Leading global product development teams. *Research-Technology Management*, **46** (6), 14–22.

Belliveau, P., Griffin, A., and Somermeyer, S. (eds) (2002) *The PDMA Handbook for New Product Development*, John Wiley & Sons, Inc., New York.

de Brentani, U. and Kleinschmidt, E.K. (2004) Corporate culture and commitment: impact on performance of international new product development programs. *Journal of Product Innovation Management*, **21** (5), 309.

Clark, K.B. and Fujimoto, T. (1990) The power of product integrity. *Harvard Business Review*, **68**, 107–118.

Clark, K.B. and Fujimoto, T. (1991) *Product Development Performance: Strategy, Organization, and Management in the World Auto Industry*, Harvard Business School Press, Cambridge.

Cooper, R.G. (2009) The state of product development. *Research-Technology Management*, **52** (1), 6.

Crawford, M. and Di Benedetto, A. (2006) *New Products Management*, 8th edn, McGraw-Hill Irwin, New York.

Duarte, D. and Snyder, N. (1997) From experience: facilitating global organizational learning in product development at Whirlpool Corporation. *Journal of Product Innovation Management*, **14** (1), 48–55.

Eppinger, S.D. and Chitkara, A.R. (2006) The new practice of global product development. *MIT Sloan Management Review*, **47** (4), 22.

Foster, R. (1986) *Innovation: The Attacker's Advantage*, Summit Books, New York.

Friedman, T.L. (2005) *The World is Flat*, Farrar, Straus and Giroux, New York.

Garrett, T.C., Buisson, D.H., and Yap, C.M. (2006) National culture and R&D and marketing integration mechanisms in new product development: a cross-cultural study between Singapore and New Zealand. *Industrial Marketing Management*, **35** (3), 293.

Gillwald, K. (1994) Cultural factors in successful innovations: culture and technical innovation: a cross-cultural analysis and policy recommendations by Horst Albach. *Futures*, **26** (8), 871.

Golder, P.N. (2000) Insights from senior executives about innovation in international markets. *Journal of Product Innovation Management*, **17**, 326–340.

Graber, D.R. (1996) How to manage a global product development process. *Industrial Marketing Management*, **25** (6), 483–489.

Hofstede, G. (1980) *Culture's Consequences: International Differences in Work-related Values*, Sage Publications, Beverly Hills.

Hofstede, G. (1994) Cultural constraints in management theories. *International Review of Strategic Management*, **5** (1), 27–49.

Jones, G.K. and Davis, H.J. (2000) National culture and innovation: implications for locating global R&D operations. *Management International Review*, **40** (1), 11–39.

Jusko, J. (2008) Product development solutions. *Industry Week*, **257** (12), 62.

Kleinschmidt, E.J., de Brentani, U., and Salomo, S. (2007) Performance of global new product development programs: a resource-based view. *Journal of Product Innovation Management*, **24** (5), 419.

McDonough, E. III, Kahn, K.B., and Barczak, G. (2001) An investigation of the use of global, virtual, and colocated new product development teams. *Journal of Product Innovation Management*, **18**, 110–120.

Nakata, C. and Sivakumar, K. (1996) National culture and new product development: an integrative review. *Journal of Marketing*, **60**, 61–72.

Roy, S. and Sivakumar, K. (in press) Managing intellectual property in global outsourcing for innovation generation. *Journal of Product Innovation Management*.

Roy, S., Sivakumar, K., and Wilkinson, I. (2004) Innovation generation in supply chain relationships: a conceptual model and research propositions. *Journal of the Academy of Marketing Science*, **32** (1), 61–79.

Sethi, R., Pant, S., and Sethi, A. (2003) Web-based product development systems integration and new product outcomes: a conceptual framework. *Journal of Product Innovation Management*, **20**, 37–56.

Sivakumar, K. and Nakata, C. (2003) Designing global new product teams: optimizing the effects of national culture on new product development. *International Marketing Review*, **20** (4), 397–445.

Smeds, R., Olivari, P., and Corso, M. (2001) Continuous learning in global product development: a cross-cultural comparison. *International Journal of Technology Management*, **22** (4), 373.

Söderquist, K.E. (2006) Organizing knowledge management and dissemination in new product development: lessons from 12 global corporations. *Long Range Planning, London*, **39** (5), 497.

Solomon, C. (1995) Global teams: the ultimate collaboration. *Personnel Journal*, **74**, 49–50.

Song, X.M. and Parry, M. (1997) Teamwork barriers in Japanese high-technology firms: the socio-cultural differences between R&D and marketing managers. *Journal of Product Innovation Management*, **14**, 356–367.

Subramaniam, M., Rosenthal, S.R., and Hatten, K.J. (1998) Global new product development processes: preliminary findings and research propositions. *Journal of Management Studies*, **35** (6), 773.

Tushman, M.L. and Anderson, P. (1986) Technological discontinuities and organizational environments. *Administrative Science Quarterly*, **31**, 439–465.

Wren, B.M., Souder, W.E., and Berkowitz, D. (2000) Market orientation and new product development in global industrial firms. *Industrial Marketing Management*, **29** (6), 601.

Zaltman, G., Duncan, R., and Holbek, J. (1973) *Innovations and Organizations*, John Wiley & Sons, Inc., New York.

pretest market models

C. Anthony Di Benedetto

A consumer-goods marketer, seeking to commercialize a new product, needs to gather market information to support the launch decision, and also to fine-tune the marketing mix (*see* LAUNCH STRATEGIES). In the past, a product manager may have tried a test market (*see also* PRODUCT TESTING), in which the product is sold in real stores under real conditions over a length of time, allowing adjustments on several components of the marketing mix. Full-scale test markets are, however, costly and time consuming, and also carry their own risks. Competitors can monitor test markets and launch their own version of the product and have even been known to sabotage test markets (by, for example, dropping valuable coupons of their own brands into the test market area). While managers still occasionally conduct a lengthy, full-scale test market for new products, they typically start with a research method that provides good quality information on customer response in a much shorter time and at much less cost. This allows the manager to screen out likely failures, before investing in the costlier test market or other market testing procedures.

One of the most common forms of market test used at this stage is the pretest market (PTM), or simulated test market. As the name implies, PTMs attempt to simulate the customer's purchase behavior in a controlled setting. The cost and time requirement is substantially lower than for a full-scale test market, and the manager avoids revealing early information about new launches to competitors.

How Pretest Markets Work

Managers were seeking alternatives to test markets years ago, and early work on PTMs was initiated as far back as the 1970s by market research firms such as SAMI/Burke (now part of ACNielsen) and BBDO. These, and other well-known PTMs such as ASSESSOR, all have at their core the awareness–trial–repeat or ATR model. In other words, the forecast long-term market share of a new consumer product is the product of its ultimate cumulative trial rate and its repeat purchase rate. In its simplest form, if a cereal producer can get 25% of the market to try a new brand, and of those, 5% become repeat purchasers, the long-run market share is projected to be 25% $\times$ 5% = 1.25%. This, of course, assumes that all potential customers are aware of the new brand, so this would have to be adjusted for the level of awareness the brand would achieve with expected advertising expenditures. The advertising agency working on the account may provide estimates on awareness, while other factors such as availability, prices, and costs may be provided by the product manager.

In addition to time and cost savings, another major advantage of PTMs is the level of confidentiality obtained by the firm. Typically about 300–600 customers may take part, and although procedures differ, most PTMs involve exposing the customer to the product (and perhaps pricing information and advertising), assessing likely trial rates, providing the customer with a sample to try at home, and following up later to determine likelihood repeat purchase. PTMs are commercially available for use by consumer-goods manufacturers, and typically provide customer information within about 14 weeks, with costs in the range $50 000–300 000 or more, depending on complexity (Crawford and Di Benedetto, 2008, p. 436).

The PTM not only provides information on the likely success of a new product but also helps diagnose where the trouble areas are. For example, if a market share is predicted to be low, it may be due to low trial and/or low repeat. A PTM will help management identify the culprit, and then take action. If trial is projected to be low, it may require that advertising copy or pricing levels be adjusted; if repeat rate is low, management may focus on why customers are so quick to switch away from the brand.

Pretest Market Procedures

Several PTM suppliers are available: two are profiled here (ASSESSOR and BASES) because these are prominent examples of each of two broad categories of PTM procedures. Both categories begin with customer trial and repeat data, but differ in the way these data are used to derive sales forecasts. ASSESSOR is typical of PTMs

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

that use mathematical models to generate the forecast, while BASES is an example of a PTM that uses a heuristic approach (a comparison against previous similar products). An in-depth discussion of both ASSESSOR and BASES is provided by Dolan (1992). Several other PTMs exist, and a good description of some of the classic ones are provided by Shocker and Hall (1986) and Urban and Katz (1983).

The ASSESSOR model[1]. The ASSESSOR model, originally developed by Silk and Urban (1978), seeks to simulate the new-product introduction process in a compressed time period (typically, well under six months). Suppose a well-known soup company is planning a new brand launch. The procedure may begin with a mall intercept or a series of phone calls, in which respondents are asked to participate in a research study. After a brief interview to determine demographics and marketing data such as category usage status (i.e., heavy, light, or no soup consumption), respondents enter a research facility and are exposed to print or broadcast advertising, for the new soup as well as several other products. The advertisements may even be placed within the context of a new TV pilot program that is itself also being tested by the network. (Procedures vary here, depending on the research agency.)

Next, the respondent moves to a mock store aisle, is provided with some cash and asked to go shopping. The new soup is, of course, one of the products in the store aisle, and the research agency may take this opportunity to try different price points or shelf locations for the soup being tested. The respondent can purchase whatever he or she wishes, and can keep the remaining cash. If the respondent did not buy the soup, then a sample of it is given as a "thank-you" gift; in any event, all participants walk away with the product in hand. A sample of respondents may be asked to stay a few more minutes, to respond to a questionnaire regarding product usage, attitudes toward the brand, and so forth. This procedure provides management with a quick assessment of likely trial, and how trial may be affected by price levels, advertising exposure, and so on.

There are variations on the above procedure. One research agency, for example, replaces the mock store with a series of purchase intention questions, and simply would give the soup to all respondents who express an interest in it.

To get a measure of likely repeat, the research agency contacts the respondents several weeks later to see how well they liked the soup. The wait time depends on the product category; if it is an item that is consumed relatively slower (like toothpaste), the agency may wait longer to call. The researcher assesses customer's product usage, positive or negative reactions, and intentions to purchase the soup in the future. The repeat rate is adjusted on the basis of this postuse data, and accounts for the possibility that customers will switch away from this brand, or switch from a competitor to this brand, in the future. To be precise, then, the repeat purchase rate here is actually an estimate of share of subsequent purchases.

Assessed trial and repeat rates are then multiplied to obtain a raw equilibrium market share. This raw market share is then adjusted down to account for projected levels of awareness and availability, to arrive at the final ultimate market share forecast. The ASSESSOR model contains a validity check on this forecast, which involves measuring customers' preferences among the different brands of soup.

ASSESSOR tends to predict market shares quite accurately. According to Urban and Katz (1983), actual performance was compared to ASSESSOR predictions, and the predictive error in the forecasts was 21.5%. However, many of the gaps between predicted and actual market share were because of changed conditions (for example, anticipated levels of awareness and distribution were never reached). After adjusting for these discrepancies, the predictive error declined to only 11.6%.

The BASES model. BASES, originally developed by SAMI/Burke, is another popular supplier of PTMs. BASES has many distinguished consumer-goods business units among its clients, including Colgate-Palmolive, Frito-Lay, Gillette, Procter & Gamble, and many more (Dolan, 1992). The BASES procedure is not too different from that outlined above for ASSESSOR. Usually, respondents are obtained through mall intercepts, shown a new product concept, and asked questions regarding their likes and dislikes regarding the

product, perceived value, and likelihood of purchase. Respondents who show an interest in the product are allowed to take some home with them, and later phone calls assess satisfaction levels with the product as well as likelihoods of repeat purchase. Unlike ASSESSOR, however, which uses mathematical modeling to convert this information to market share projections, BASES uses a heuristic approach. Very simply, BASES has a database of thousands of prior PTM studies. They select data from comparable previous product introductions, and use these data to calibrate raw trial and repeat scores. These adjusted data are then run through an ATR model to develop the sales forecast. Another heuristic model, NEWS, developed by the advertising agency BBDO, does not rely on data from previous product introduction, but rather uses survey data derived only from the brand under study (Pringle, Wilson, and Brody, 1982).

Like ASSESSOR, BASES has established a good forecasting track record. According to Dolan (1992), Burke claims that 90% of their sales, trial, and repeat forecasts are within 20% of the actual levels, while over half are within 10%.

New Trends in Pretest Markets

We are now starting to see applications of PTM models in the forecasting of consumer durables and business-to-business products. General Motors, for example, has been using a new measurement method known as *information acceleration* (IA) to assess purchase intentions of new cars (Hauser, Urban, and Roberts, 1990). In IA, respondents enter a virtual buying environment that simulates, in this case, the car purchase situation. They interact with virtual salespeople and hear word of mouth from virtual customers, they see print and broadcast ads, and can even virtually walk around the car showroom to look at computer prototypes of new models.

Consumer goods can also be virtually tested, as in the visionary shopper (VS) procedure. In VS, the respondent enters a virtual store and can "shop," read labels and price information, and make purchases, all by touching the computer monitor. Companies are increasingly using VS or similar virtual-testing procedures as an important part of the PTM process (Burke, 1996; Rosenberger and de Chernatony, 1995).

Criticisms of Pretest Markets

PTMs have received their share of criticism. The mall intercept or related techniques essentially create a laboratory experiment environment (that is, people know they are being tested). In addition, the mock store aisle is clearly not the real thing, and the advertising stimuli are presented in what some say is an unrealistic way. It is possible that these false conditions make the results suspect. PTM suppliers argue, however, that few other methods can provide good quality forecasts in such a short time. Nevertheless, there still are drawbacks. We saw earlier that the forecasts will be only as good as the projections of awareness and availability (as well as other projections) that come from management; another problem is that only final consumer response is considered. Our soup manufacturer might find that the new soup brand scores high in PTMs with consumers, but unfortunately is not liked by retailers (maybe the newly designed cans are hard to stack and fall off the shelf?) and therefore they refuse to carry or promote it. In addition, PTMs might have less value for new-to-the-world products, or for products that are not sold via personal selling or point-of-purchase promotions (Crawford and Di Benedetto, 2008, p. 439).

Endnotes

[1] The sections on ASSESSOR and BASES derive from Crawford and Di Benedetto (2008), pp. 436–439 and Dolan (1992).

Bibliography

Burke, R.R. (1996) Virtual shopping: breakthrough in marketing research. *Harvard Business Review*, **74** (2), 120–131.

Crawford, M. and Di Benedetto, A. (2008) *New Products Management*, 9th edn, McGraw-Hill/Irwin, Burr Ridge.

Dolan, R.J. (1992) *Researching and Monitoring Consumer Markets*, Harvard Business School Publication No. 9-592-088, Harvard Business School Press, Boston.

Hauser, J., Urban, G.L., and Roberts, J. (1990) Prelaunch forecasting of automobiles. *Management Science*, **36** (4), 401–420.

Pringle, L.G., Wilson, R.D., and Brody, E.I. (1982) NEWS: a decision-oriented model for new product analysis and forecasting. *Marketing Science*, **1** (1), 1–30.

Rosenberger, P.J. and de Chernatony, L. (1995) Virtual reality techniques in NPD research. *Journal of the Market Research Society*, **37** (4), 345–355.

Shocker, A.D. and Hall, W.G. (1986) Pretest market models: a critical evaluation. *Journal of Product Innovation Management*, **3** (3), 86–107.

Silk, A.J. and Urban, G.L. (1978) Pre-test-market evaluation of new packaged goods: a model and measurement methodology. *Journal of Marketing Research*, **15** (2), 171–191.

Urban, G.L. and Katz, G.M. (1983) Pre-test-market models: validation and managerial implications. *Journal of Marketing Research*, **20** (3), 221–234.

idea management

Jeffrey Baumgartner

Introduction

Idea management is a structured business process for capturing focused business ideas from a population, typically the employees of a company, and evaluating those ideas to identify the most promising ones. The process is typically facilitated by an enterprise software application with a front end for the submission of ideas and a back end where those ideas are reviewed. The front end, usually a set of web pages, is open to a large group that may consist of some or all employees of a company as well as, in some instances, business partners, customers, stakeholders, and others. This group typically numbers several hundred people, although some applications permit thousands of users. The back end normally may only be accessed by managers who will use the evaluation tools to determine which ideas should be promoted for implementation.

In addition to the software, substantial human intervention from the idea management team – in other words, the person or people overseeing the process in a given organization – is critical to an effective idea management initiative. Promotion of innovation and the idea management process, senior management buy-in, a good communication plan, and a rewards scheme are among the actions that help make an idea management initiative a success. And, although it may seem obvious, the final step of an effective idea management initiative is the profitable implementation of the best ideas.

Although idea management is considered an innovation process, it is arguably more of a CREATIVITY tool. Idea generation is a creative process and, as we shall see, most viable idea management systems are based on creative problem solving (CPS) methodology.

If idea management sounds a lot like a modern suggestion box, that is not surprising. Not only is the process arguably derived from suggestion schemes but many firms still use a variation on the suggestion box as the basis for their idea management process. However, as we shall see, this is not usually an effective approach.

Suggestion Schemes

The predecessor to idea management was the suggestion box, which is essentially the front end of a suggestion scheme. A suggestion scheme is a simply structured approach to capturing ideas. A suggestion box, or other repository, is set in a public space within an organization. Employees are invited to submit suggestions. Periodically, management reviews the submitted suggestions and decides on which ideas to implement.

The first recorded use of a suggestion scheme was by the Japanese Shogun in 1721, who offered awards for ideas (Beddows, 2001). The British Navy implemented a suggestion scheme in 1770 as a means for allowing all their personnel to offer suggestions without fear of consequences. At that time, any suggestion that contradicted the thinking of a senior officer was subject to reprimand and, in extreme cases, hanging (Robinson and Stern, 1998)! The first suggestion box and structured suggestion scheme were implemented in 1880 at the Scottish shipbuilder William Denny and Brothers (Beddows, 2001).

Suggestion schemes have persevered over the centuries, with many companies providing them for employees as well as customers. With the dawn of the web in the early 1990s, and its massive take-up in the following decade, many organizations have implemented e-mail and web-based suggestion schemes. In the former, an e-mail address is set aside for the collection of ideas and employees are encouraged to use it. In the case of web-based systems, employees complete an on-line form to submit their ideas to a database. In knowledge-based companies, where most employees have computers on their desks, such systems make it easy to submit ideas.

Most suggestion schemes have more clearly thought-out front ends than back ends. Indeed, the simplest, ad hoc systems essentially have no back end. Like the suggestion box of old, an employee – often a manager – regularly looks at every submitted idea and makes a decision as to whether to escalate each idea – in other words, send it for further review such as the development of a business case – or to dispose it of. More sophisticated systems boast a range of tools that allow collaboration on ideas, provide structured review processes, provide voting, and much more. In many cases, such sophisticated

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

tools are called "*idea management software*". However, the lack of a means to focus ideas arguably prevents a suggestion scheme from actually being considered an idea management solution.

Fatal Flaw with Suggestion Schemes

In spite of their popularity, suggestion schemes in medium to large organizations suffer from a serious flaw: they typically fail after 12–18 months, in this author's experience. The cause of failure is almost inevitably the lack of focus of suggestion schemes, which leads to several undesirable consequences:

1. If the suggestion scheme is well promoted in the organization, it leads to a high number of submitted ideas. A huge influx of such ideas is often too much for a single person or team to manage. As a result, they are unable to process the ideas in reasonable time and provide feedback to the idea submitters. Idea submitters, in turn, get no information about their submissions and assume management is not really interested in their ideas. Hence, they do not continue to submit new ideas.
2. If the system is blind, that is, idea submitters cannot view each other's ideas, there tend to be a very large number of duplicate and very similar ideas. This not only slows down the processing of ideas, but poses problems to the suggestion scheme's manager, specifically as to who to acknowledge as the idea initiator and how to prevent hard feelings among others who submitted similar ideas and who are not recognized as the initiator.
3. Many submitted ideas are not relevant to business needs. Hence they need to be rejected. This results in a high failure rate, which is demotivating to employees and tends to discourage continued participation in the scheme.

To illustrate the last point, the author is familiar with a multinational manufacturing and engineering firm, which launched a suggestion scheme in its Chinese head office. To ensure the open communication and fast flow of ideas, the team in charge of the scheme was required to review every idea as soon as it came in and give feedback to the idea owners within a week. However, most of the ideas were not relevant to business needs and had to be rejected. While most suggestion schemes fail within 18 months, this one lasted far less than a year. In retrospect, it is not difficult to understand why. Employees submitted ideas that they felt good about. Within a week, they received a critical report rejecting their ideas. Essentially, the company was communicating that it did not like the employees' ideas. This became particularly apparent as employees discussed results with each other. The problem was exacerbated by Chinese culture. Employees felt they had lost face by being rejected by management! As a result, an expensive suggestion scheme not only resulted in very few viable suggestions but created negative feelings among employees, especially the most creative thinkers.

In 2007 and 2008, several large organizations launched public, web-based suggestion schemes in which anyone could to register and submit ideas. Although these have succeeded as publicity tools and communication tools – facilitating customer feedback – they have not demonstrated impressive results as innovation tools. A quick look at the most popular of these schemes will reveal that thousands of ideas, including numerous duplicates, have been submitted and that at most a few tens have been reported as being considered for implementation. Moreover, if one assumes that each idea will take an employee an average of 10 minutes to review, each 1000 ideas will require 166.7 man hours to process. That is not an efficient back end.

Fortunately, there are better and more efficient methods of capturing focused business ideas from large groups of people.

Creative Problem Solving (CPS)

Creativity experts have long known that great ideas very rarely explode out of nowhere. Rather, they are developed in the process of collaboratively seeking a solution to a problem. Although creative problem solving has been around since humans attained the ability for abstract thinking, the process was probably first defined by Alex Osborn, the inventor of BRAINSTORMING, and a cofounder of Barton, Durstine & Osborn,

an advertising agency that is now a part of BBDO Worldwide. He published his approach to brainstorming in 1948 in his book *Your Creative Power*. Here he laid out the notion of a defined problem and an open, uninhibited period of idea generation (Osborn, 1948). Over the next decade, Mr. Osborn together with Dr. Sidney Parnes developed the brainstorming approach into a more structured process they called *creative problem solving* or *CPS*. Since then, CPS has been further refined and developed, largely at the International Center for Studies in Creativity at Buffalo State College (Firestien, 1996). However, many variations of the process have been adopted by practitioners of creativity and innovation. For the purposes of this article, we will consider a simple seven step process:

1. identifying a problem, a goal or a need;
2. researching the problem to better understand your actual requirements;
3. formulating a creative challenge;
4. generating ideas in response to the challenge in an open, uninhibited and noncritical environment;
5. evaluating and combining ideas to develop viable solutions;
6. drawing up an action plan for implementing the ideas;
7. implementing the ideas.

Applying CPS to the suggestion scheme results in what idea management experts variously call an *ideas campaign*, an *event* or a *challenge*. We will use the term *ideas campaign* in this article. It is such an effective approach that it forms the basis of most of the established commercial idea management software products available today. Let us look at how an ideas campaign works.

Ideas Campaigns

An ideas campaign essentially follows the CPS approach with minor modifications. In particular, the idea generation phase of a CPS event, such as brainstorming, typically lasts a few hours or at most a day or two and includes a relatively small group of participants. In an ideas campaign, in which hundreds or thousands of people in various locations participate, the idea generation phase typically lasts anywhere from one to four weeks – sometimes even longer. Idea management software backed by a robust, scalable database facilitates participation by such a large number of people.

Of course, an ideas campaign is a single event. However, by running simultaneous and sequential ideas campaigns, an organization will have a sustainable, effective idea management process that focuses innovation on business needs. We look at the advantages of the ideas campaign approach to idea management over suggestion schemes later in the text. First, let us take a look at the elements of an ideas campaign.

Innovation challenge. An ideas campaign starts with an innovative challenge, which is typically a terse, focused question such as "What new products might we develop?" or "What business opportunities might we exploit in China?" As with CPS, this challenge should not be hastily formed. Rather it should be developed through internal questioning and research (VanGundy, 2007). In this author's experience, too many managers make their challenges overly vague – using only a word or two, for example: "Marketing ideas" – or overly detailed, bringing in market research, analysis, and more. In the former case, participants are unsure of the managers, needs and in the latter case, the excess of details tends to discourage participants as they fear they will submit an unsuitable idea.

Idea generation. Once the ideas campaign is launched with an innovation challenge, participants are invited to submit their ideas. In an ideal idea management system, all submitted ideas are visible to all participants and some form of commenting function permits collaborative idea development. During the idea generation phase, it is important that management be supportive and encourage submission of ideas. Criticism or rejection of ideas at this stage at best discourages creativity in idea submission and at worst discourages participation altogether. As Mr. Osborn found out, while developing brainstorming, criticism during the idea generation phase makes people fear ridicule, personal criticism, and possibly even reprimand. As a result, they will be more cautious – and hence less

creative – in their idea submission or may opt not to submit any ideas at all. From the participants' point of view, if idea submission results in criticism, not submitting ideas is a good way to avoid that criticism.

Evaluation and review. At the end of the idea submission phase, it is time to identify the best ideas. This is done via evaluation of ideas and combining complementary ideas together. There are numerous idea review approaches used in different systems. Some of the more commonplace methods are described below.

Evaluation matrix. The evaluation matrix is essentially a peer review process. It compares an idea against a set of criteria. Evaluators indicate on an online form how well they believe the idea meets each criterion. The software then calculates and compiles an overall score and generates a report. Evaluators are typically internal experts with substantial knowledge relevant to the theme of the innovation challenge. Having multiple experts ensures a more balanced review than would be the case if an individual were to evaluate the ideas alone.

Pass/Fail. A simple pass/fail selection is often necessary to hone down a large number of ideas into a small, more manageable number of ideas for more detailed evaluation. A pass/fail selection should be based on one or more criteria to make the selection process as objective as possible.

Voting. Users of the system vote for ideas. The ideas with the most votes are selected for further processing. Although popular, voting is not normally a scientific approach to idea review. Highly innovative ideas are seldom the most popular. Indeed, they often seem the most absurd initially. As a result, they are unlikely to receive many votes. Incrementally innovative ideas, on the other hand, are easier to grasp and less radical. Hence usually they do better in open voting systems. In other words, putting ideas to a vote typically results in less innovative ideas getting high scores and highly innovative ideas getting low scores.

SWOT Analysis. A strength, weakness, opportunities, threats (SWOT) analysis is a widely used business method for evaluating ideas, particularly product and marketing ideas. As a result, it is also a useful tool in an idea management process that includes generating such ideas.

And beyond. Once ideas pass the initial evaluation tests, they are often developed into business cases, turned into prototypes, subjected to market analysis, and more, before businesses invest in their implementation. This is particularly true where ideas require a substantial investment to launch. However, these more advanced tools are seldom part of an idea management tool or process. Product development and project development are usually considered entirely different disciplines from idea management and, as a result, are overseen by different teams.

Advantages of Ideas Campaigns Over Suggestion Schemes

Idea management by ideas campaign clearly has several advantages over a suggestion scheme.

1. Creative idea generation is focused on actual business needs as identified by managers who are typically decision makers. As a result, ideas are more likely to be relevant and useful. This means there are more suitable ideas and fewer rejected ideas.
2. All ideas focus on solving a particular problem. This facilitates combining ideas to build better ideas and allows for comparative evaluation. In other words, rather than just reviewing one idea for viability, managers can review numerous solutions to identify the best idea or ideas.
3. A creative challenge is far more motivating than a simple request for ideas. Compare sitting in a brainstorming session where the facilitator simply says "Give me your ideas!" against a session where the facilitator asks "What creative uses might you make of a brick?" It would clearly be easier to generate ideas for the latter! The same is true for idea management.

Indeed, the advantages of the ideas campaign approach, over the suggestion scheme, as a

process for generating viable business ideas is so clear that it has become the basis of most idea management systems. Nearly all commercially available idea management software products use a variant on this approach and consultants also use it as the structure of any idea management process they implement. Nevertheless, there are a few firms that still use suggestion schemes, and many organizations wishing to design their own idea management software almost inevitably start with a suggestion scheme. This is typically because the system is designed by software programmers rather than innovation experts.

Putting it All Together

Most sustainable idea management initiatives today are based around running regular ideas campaigns. Often ideas campaigns feed each other. For instance, a manager runs an ideas campaign for new-product ideas. One of the ideas is selected for implementation. As a result further ideas campaigns are run about how to implement the product's features, how to package the new product, how to position the product in the marketplace, and so on.

In addition to ideas campaigns, most idea management processes also provide a means of submitting spontaneous ideas that do not fit in any running ideas campaigns. This is a good means of capturing good ideas that may come to employees as a result of outside stimuli. However, care must be taken not to become reliant on spontaneous ideas as the focus of an idea management program. Otherwise it becomes little more than a suggestion box.

For the most part, idea management programs based on ideas campaigns use commercially available idea management software to facilitate the process.

In addition to software and structure, there are several human components that need to be added to any successful idea management initiative.

Factors in the Success of Idea Management Initiatives

Even with the best software and operational plan, many idea management initiatives fail to meet expectations largely owing to poor levels of participation. In this author's experience there are two reasons for this. The first is that people simply do not know about the initiative. The second, more commonly, is that people claim they have no time to participate. Indeed, at least one study has identified "lack of time" as the number one hurdle to organizational innovation (Morris, 2005). Yet a moment's thought is sufficient to realize that this makes no sense. Every employee in an organization has at least 35 hours per week of time. Most have many more hours. Hence people are really saying that participating in innovation initiatives is not of sufficient importance for them to devote some of their time, and that it is of very little priority. Moreover, if most people in an organization feel that innovation is not a priority, there is little likelihood that idea management or any other innovation activity will achieve its aims.

For any idea management project to succeed, it is important to make innovation a high-level priority in the minds of the employees expected to participate. This achieved, management can work toward continually improving the results of their ideas campaigns. In this author's experience, management buy-in, a good promotional plan, a rewards scheme, and a willingness to take on the risk of implementing creative ideas are the key ingredients to a successful idea management initiative.

Management buy-in. If top management is not behind an idea management initiative, they cannot expect subordinates to bother with it either. If employees see that management is not demonstrably interested in participating in idea management, they will quickly learn that there is no personal benefit in participating themselves. On the other hand, if managers are actively involved in idea management, submitting ideas themselves, and personally thanking submitters of ideas for their contributions, employees will see that management is serious about innovation and that it is a corporate priority.

For managers wanting to introduce idea management, this is critical. They must not only look at the software and resource requirements. They must be prepared to invest their own time and energy in it. Innovation and idea management must be made a top priority.

Promotional plan. If employees do not know

about their firm's idea management activities, they can hardly participate. If they do not understand the importance of the idea management initiative, they are unlikely to prioritize participation. Hence, any idea management process planning needs to include a promotional plan designed to communicate the details of the process, why the idea management initiative has been launched, why everyone should participate, the benefits to the organization, and the benefits to the participants themselves.

The promotional plan should include an explanation of the process, information about the rewards system (more below), and a description of how to participate, including a URL of the idea management software, log-in procedures, and anything else people will need to know to use the system.

Beyond promoting the idea management initiative to employees, management should also promote it to clients, stakeholders, business partners, and possibly even the general public. This helps clarify in the employees' minds the importance of the innovation initiative and tells the outside world that the firm is very serious about innovation.

Rewards scheme. Rewards, for submitting ideas to an idea management system, are a good way to motivate people to participate. However, a poorly designed rewards scheme can actually discourage people from participating. So any rewards scheme needs to be carefully designed and should follow two basic rules.

Lots of small rewards. It may seem counterintuitive, but expensive rewards such as large cash prizes or highly valuable gifts can actually destroy an idea management process. When rewards are too big, people tend to begin acting selfishly. Managers steal subordinates' ideas, people do not collaborate – which minimizes the development of ideas – and people act solely to win rewards rather than to innovate on behalf of the company.

Small rewards, on the other hand, are not valuable enough to encourage selfish action. Rather, the rewards are appreciated by employees as acknowledgement by management of each person's participation. Moreover, people can be competitive with small rewards without causing excessive bad feeling from those who have received fewer rewards.

In the early stages of an idea management implementation, a small reward such as points that can be exchanged for gifts, a chocolate, or a piece of fruit for every idea is an excellent means of encouraging employees to log into the idea management system and submit ideas. Once employees have done this, and provided the software is sufficiently intuitive to use, employees will quickly become used to using it and participating in the initiative.

Reward for quantity or creativity – but not for quality. A frequently made mistake in rewards schemes is to offer rewards for "the best ideas." However, research has shown that this actually results in reduced creativity and fewer ideas (Sawyer, 2007)! But with a moment's thought the reason for this is clear. When participants are told to produce good ideas or compete to submit the best idea, they aim to please management. As a result, they submit ideas they believe managers will like rather than risk submitting an unusual idea that management might not like. Ideas that management are perceived to like will be ideas that are similar to ideas and actions management has taken in the past. In other words, unusual and hence creative ideas are rejected in favor of ideas that are at best incremental improvements on existing concepts and hence not very creative.

Rewards for the most creative ideas or the most outrageous ideas, on the other hand, can be an effective means of encouraging creative thinking. This indicates to participants that they are invited to open their minds and think freely. The result tends to be fewer ideas than in ideas campaigns with quantity-based rewards or no rewards, but those ideas do demonstrate a higher level of creativity (Sawyer, 2007). Thus such rewards can be an effective strategy once the idea management process is operational and familiar to employees.

Group rewards versus individual rewards. One last note – in general, rewarding teams and groups is more effective than rewarding individuals, particularly if the rewards are larger. This has consistently been demonstrated in research and is intuitive. The

advantage to idea management, like other group idea generation methodologies, is that people can collaborate together to build bigger ideas than would be possible for any team member as an individual to achieve. However, if rewards go to individuals, team members have an incentive not to collaborate as doing so might result in a colleague being rewarded for the idea originator's solution. On the other hand, if an entire team, group, or business unit is rewarded, all members have an incentive to collaborate and cooperate. This results in better developed, more innovative ideas.

Willingness to take on risk. No matter how well organized an idea management initiative is, it will not be an effective innovation tool if management does not implement the ideas it generates. As obvious as this notion is, it is surprisingly frequently not followed.

Potentially innovative ideas are inherently risky. They are new, untested, and deviate from the norm. The more innovative an idea is, the greater is the risk of failure. Conservative managers are often unwilling to take on such risk and ironically fail to implement the most potentially innovative ideas their people develop.

As a result, any idea management initiative should be accompanied by a risk management structure designed to accommodate innovative ideas. This might include a specified budget for high-risk ideas as well as a strategy for risk management. Such a strategy should include milestones against which an idea's implementation may be measured. If the idea fails to meet or exceed milestones, it should be reviewed. If the problems cannot be solved, the idea must be dropped, and another innovative idea implemented. Such an approach reduces risk and yet ensures a steady stream of innovative ideas are being tested and implemented. And that, more than anything, demonstrates the successful implementation of an idea management system.

Bibliography

Beddows, A. (2001) Suggestions Schemes for the Future, Management Service, UK, February.

Firestien, R.L. (1996) *Leading on the Creative Edge*, Pinon Press, Colorado Springs.

Miller, B., Vehar, J., and Firestien, R. (2001) *Creativity Unbound*, Innovation Resources, Williamsville.

Morris, W. (2005) A Survey of Organisational Creativity, http: //www. jpb. com/ creative/ Organisational CreativityMorris.pdf (accessed 17 February 2009).

Osborn, A.F. (1948) *Your Creative Power*, Charles Scribner's Sons, New York.

Robinson, A. and Stern, S. (1998) *Corporate Creativity*, Berrett-Koehler, San Francisco.

Sawyer, K. (2007) *Group Genius*, Basic Books, New York.

VanGundy, A.B. (2007) *Getting to Innovation*, Amacom, New York.

open innovation

Joel West

Introduction

Control of research and development has been long assumed to be crucial to the success of large firms in technology-intensive industries. The rise of the great vertically integrated chemical, automobile, and electronics companies in the early twentieth century – and IBM in the postwar era – have been held out as examples of how industrial R&D enabled firms to offer new products and capture new markets, as portrayed by the business histories of Alfred D. Chandler, Jr. studying large industrial firms such as BASF, General Motors, and Unilever. In the United States, the industrial research of IBM and AT&T's Bell Labs was exemplars of corporate R&D that spawned firm and societal success.

However, such exemplars had important limitations. Henry Chesbrough studied the successful and unsuccessful efforts of Xerox to commercialize innovations from its famed Palo Alto Research Laboratory (Xerox PARC), concluding that it was essential for firms to align their innovation efforts to business models (Chesbrough and Rosenbloom, 2002).

In his book *Open Innovation*, Chesbrough (2003a) concluded that the traditional vertically integrated model was failing as firms increasingly paid a penalty for failing to commercialize their innovations, because of greater customer demands and market pressures. Meanwhile, knowledge is increasingly dispersed and mobile, as researchers who helped develop a technology can commercialize it by moving to a new employer or start their own company.

Chesbrough's conception of open innovation has important antecedents in both research and practice. Prior research includes the work on profiting from innovation by Teece (1986) and Eric von Hippel's prolific work on user innovation. Extant managerial practice includes product modularity, component-based innovation, and systems integration – particularly in electronics-based industries – as well as patent licensing business models from companies like Dolby Labs, Qualcomm, and dozens of start-up biotech companies.

Nonetheless, Chesbrough (2003a) has inspired a new subtopic within innovation studies, as researchers have studied a broad range of efforts by firms to source and commercialize innovations outside their firms' boundaries.

Definition

Chesbrough (2003a: xxiv) introduced open innovation as "a paradigm that assumes that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as the firms look to advance their technology." In a subsequent book, Chesbrough offered this expanded definition:

> Open innovation is the use of purposive inflows and outflows of knowledge to accelerate internal innovation, and expand the markets for external use of innovation, respectively. [This paradigm] assumes that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as they look to advance their technology (Chesbrough, 2006: 1).

As an editor of the 2006 book, the author knows that Chesbrough intentionally chose "purposive" over "purposeful," although this choice forces people to go to their online dictionary. Something more concise using simpler language was required for the author's own teaching, presentations and blogging, so he developed these two alternatives:

> Open innovation means treating innovation like anything else — something that can be bought and sold on the open market, not just produced and used within the boundaries of the firm.

> Firms that embrace open innovation employ markets rather than hierarchies to obtain and commercialize innovations.

Both of these latter definitions focus on the transaction cost economics (TCE) critique that was implicit in Chesbrough's original definitions. TCE argues that managerial hierarchies (in this case, vertical integration) are only preferable when market transactions would be too

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

expensive due to the costs of search, bargaining, contracting, or enforcement.

These latter definitions also suggest that open innovation is possible only with efficient and effective markets for innovation that provide for search, valuation, and bargaining, as well as legal institutions that enable contracting and enforcement. This is consistent with Chesbrough's (2003a) original observations about the importance of intellectual property (IP) and markets for buying and selling such IP.

Even with such markets, transactions will only occur if there is a sufficient population of willing buyers and sellers participating in the market, since unrealistically low (or high) prices will deter participants from using open innovation in preference to vertical integration. As Chesbrough (2003a) notes, accurate pricing is made more difficult when a firm cannot precisely value its IP, particularly when an innovation requires the firm to develop a new business model.

That is not to say that open innovation is only possible through markets: Chesbrough repeatedly encourages firms to accept (but not create) free spillovers of innovation from external sources. Chesbrough (2003b) identifies a key category of external sources which he terms "innovation benefactors," to include those (such as the US National Science Foundation, National Institutes of Health, or disease-specific nonprofit foundations) that provide funding for innovation. At one point university research (whether internally funded or government funded) served as sources of free spillovers, but after the 1980 Bayh–Dole Act, US universities have moved to aggressively patent and monetize their innovations. Under this regime, opportunities to exploit large-scale free spillovers of government-funded university research – such as those that created the Internet – may have become a thing of the past.

Researchers have identified specific aspects within the open innovation process. All draw a distinction between external commercialization of internal innovations and internal commercialization of external innovations. (Firms practicing open innovation are assumed to balance those practices with vertically integrated commercialization where that is more appropriate.)

Chesbrough's (2003a) initial definition focused on firms finding markets for their internal innovations, while identifying and utilizing external sources of innovation. For the large firm, the key determinant for a firm commercializing its own (or others') innovation was the alignment of the innovation to new or existing business models. For small firms, the need to go outside for commercialization is consistent with the Teece's (1986) conclusion that to profit from their innovations, innovators who lack complementary assets necessary for commercialization must contract for the missing parts of the value proposition.

To Chesbrough's focus on commercializing internal and external innovations, West and Gallagher (2006) identified a third issue: motivating a supply of external innovations. While the IP market view assumes external innovations can be purchased, their studies identified the case where innovations are obtained from external sources through nonpecuniary incentives such as reputation enhancement.

Building on the internal/external distinction – which they term "inside-out" and "outside-in" – Enkel, Gassmann, and Chesbrough (2009) extended it to include a third mode for innovation creation and commercialization: the "coupled process" of cocreation with complementors and other parties. This is consistent with (as will be discussed later) community-level and industry-level open innovation processes, as well as overlapping user innovation and cumulative innovation research.

Researchers have used differing (often implicit) definitions of what constitutes an "innovation": in the original exemplar of Chesbrough (2003a), innovations bought and sold by firms included basic research, patents, and manufactured components. As in user innovation, the focus often emphasizes the openness of the process rather than the innovativeness of the outcome, thus including as "innovation" incremental improvements on well-developed technologies.

Chesbrough's use of "open innovation" is not without its controversy. "Open" refers to the permeability of an organization to its external environment, as emphasized in open systems theory by organizational theorists such as Daniel Katz, Robert Kahn, and W. Richard Scott.

However, the emphasis in open innovation on monetizing knowledge flows across organizations boundaries puts it at odds with other theories (and their associated policies and normative prescriptions) that assume that openness means knowledge sharing and nonmonetized knowledge flows. Another controversy arose when some researchers questioned the novelty of the open innovation process – although Chesbrough and others freely acknowledge the importance of prior research and practices.

Related Theories

Open innovation is not the only stream of research that considers the production of innovation across (or outside) organizational boundaries – in contrast to the Chandlerian vertically integrated ideal – nor is it the only one to consider separately the creating and diffusing (including commercialization) of innovations. The two most closely related streams are user innovation and cumulative innovation, which share a common interorganizational perspective, but differ in their assumptions, findings, and implications for policy and practice (West, 2009).

As created by Eric von Hippel, research on user innovation emphasizes the unique (and "sticky") knowledge of unmet needs held by informed users (see von Hippel, 2005 for a review and summary). This knowledge gets applied and diffused in a variety of ways. One approach is for firms to identify and enlist lead users to suggest new avenues for firms in their product development efforts (*see* LEAD USERS). A complementary stream of research describes mechanisms such as "user toolkits" that allow users to prototype and even implement solutions to their own problems. Other research has considered communities of users sharing solutions with each other, and users directly commercializing their own ideas by becoming user entrepreneurs.

Meanwhile, research on cumulative innovation (or "collective invention") tends to focus on radical innovations that are technologically immature, but are successively refined by competing producers (see West, 2009 for citations). Research on cumulative innovation embraces a range of assumptions about knowledge sharing, from the intentional collaboration of a communal group of producers to the unintended (and uncontrollable) spillovers between rivals locked in a Darwinian struggle for success.

Some of these researchers have used the term *open innovation* to refer to these more collaborative, less profit-oriented innovation processes. More recently, von Hippel (2005) consciously used the phrase "open and distributed innovation" to refer to this sort of collaborative process, as distinct from the concept of Chesbrough and his successors.

A major difference between open innovation and the other two streams is their respective assumptions (and recommendations) about intellectual property rights. Open innovation usually assumes that markets for innovation are enabled by strong appropriability through strong IPRs (intellectual property rights), particularly patents (West, 2006). Further, the other two streams warn of the consequences of excessively strong IPRs in hindering collaboration, user innovation and spillovers, and in some cases explicitly recommend relatively weak IPRs (West, 2009).

While they have distinct emphases, these three streams do have common areas of interest, most notably in open source software (*see* OPEN SOURCE). Software by its nature is a nonrivalrous information good that is easily shared (or transmitted), while large software systems have the malleability and modularity that permit distributed production of individual components. The rise of free software (and later open source software) during the 1990s and 2000s enabled distributed use and production of entire systems. There is the open question (cf. West and Lakhani, 2008) as to whether a typical open source project generates "innovation" as normally defined in the literature – particularly for projects like Linux that are freely distributed imitations of existing technologies – but the latter are included in the cost reduction definition of "radical innovation" used by O'Connor and her colleagues (*see* RADICAL INNOVATION).

The largest body of such research on open source is present in the user innovation stream (see von Hippel, 2005 for a summary of relevant work). Many of the most successful early

free or open source software projects originated with software developers "scratching an itch," mapping directly onto the user innovation paradigm of users addressing their own unmet needs. Meanwhile, users worked in peer and hierarchical communities to correct defects on the shared software product and provided mutual tech support. They also added new capabilities by using user toolkits that allowed creation of plug-in modules and other software components.

At the same time, open source software provided an economical and increasingly important source of external innovations for IT firms during this same period. West and Gallagher (2006) identified four different ways that firms were employing open innovation in their open source strategies. The first was to spin out innovations from the firm – not as Chesbrough (2003a) suggests into new firms, but into new nonprofit open source communities. The second is to use open source as a form of pooled R&D, as when major computer vendors collaborate to produce Linux for use with their respective computers. Finally, they identified two forms of software hybrids combining proprietary and open source components: those where the external open source technology was the core (while the firm sold complements), and those where open source complements increased the value of a proprietary core.

Phenomena of Interest

The open innovation paradigm was conceived based on Chesbrough's (2003a) observations of managerial practice. This included both the paradoxes faced by Xerox's Silicon Valley researchers in commercializing IT breakthroughs within a photocopier company, and IBM's wrenching transformation during the 1990s away from vertical integration toward a more agnostic attitude toward sourcing and commercializing innovations.

Both in Chesbrough's telling and subsequent research, the open innovation concept is often associated with information technology or other technology-based industries (Chesbrough and Crowther, 2006). However, the recent research of open innovation has had limited impact on Silicon Valley, perhaps because the key practices were already quite common here. Building on product modularity, Silicon Valley firms have long since built products such as PCs or network equipment by purchasing components (e.g., microprocessors and random access memories), enabling open innovation business models in which supplier firms monetize their innovations by selling standardized components to systems integrators (West, 2006) (*see* PRODUCT MODULARITY). (Figure 1).

Meanwhile, the sorts of spin-offs (and acquisitions) identified by Chesbrough and Rosenbloom (2002) to commercialize technology had become quite common in this high tech region by the time that *Open Innovation* (Chesbrough, 2003a) was published. Similarly, new biotech start-ups have for decades licensed internal innovations to large pharmaceutical firms, solving the classic problems identified by Teece (1986) that small start-ups face in profiting from technological innovation because they lack the necessary distribution, manufacturing, or support.

There is anecdotal evidence that open innovation is of greater interest – or at least has greater potential for business transformation – in more mature industries where Chandlerian integration is the norm. Chesbrough's (2003a) original study of open innovation highlighted Procter & Gamble's open innovation practices, and as a direct consequence several other consumer products companies have also sought to develop open innovation practices. Some have suggested that open innovation could become the new model for the once vertically integrated automotive industry, as twenty-first century electric car start-ups such as Fisker and Tesla concentrate on engine and integration expertise (emulating the IT model) while buying or outsourcing other activities in the value-creation process.

The diffusion of the concept to established industries has been fueled by the innovation infrastructure. External sourcing of innovation through intermediate markets had been documented prior to the publication of *Open Innovation* (see Chesbrough and Crowther, 2006 for citations). Since that time, Chesbrough has worked directly with the Industrial Research Institute to promote open innovation. The concept was also highlighted by professional organizations such as the Product Development

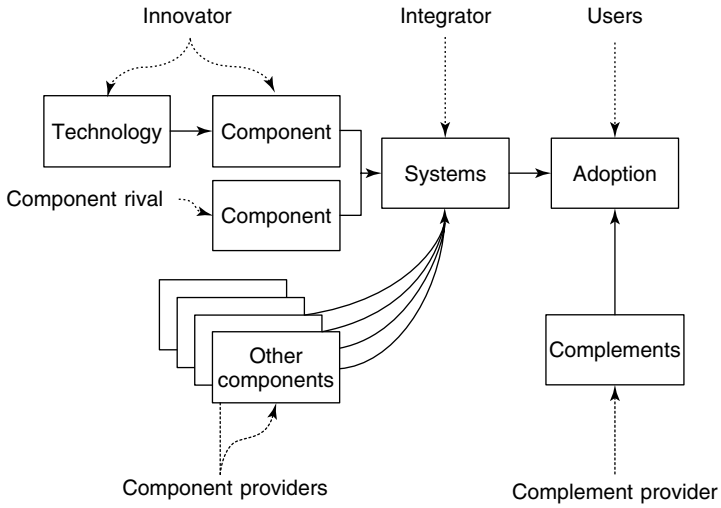


Figure 1 Open innovation for IT products, based on components and systems integration (adapted from West 2006).

& Management Association with consultants and workshops springing up to promote best practice and offer turnkey solutions.

If open innovation is gradually spreading beyond technology industries, it still seems (to use the terminology of Enkel, Gassmann and Chesbrough, 2009) to be mainly focused on the "outside-in" rather than "inside-in" (cf. Chesbrough and Crowther, 2006). This emphasis dominates the popular press, research and business practice, and is likely because of the limited number of firms generating significant quantities of unmonetized IP.

However, the "outside-in" also appears to have an inherent attraction for CEOs seeking to eliminate head count and reduce R&D budgets. For example, in 2008 Hewlett-Packard – the oldest and largest Silicon Valley tech company – renamed its "University Relations" office as "Open Innovation," to focus on developing and sourcing external technologies from universities and small firms in the face of ongoing layoffs in its research labs. This inflow of external innovation also applies to large companies seeking to continue innovation long after declining margins cause deep cuts in R&D, as evidenced in the embrace of open innovation by former telecommunications monopoly carriers such as BT and Deutsche Telekom.

Researchers have also been studying the increasingly clever practice of stimulating, identifying, and procuring external innovations. These range from corporate innovation grants provided by HP in the range of $25 000–100 000 to the $200 cash grants given by Threadless for winning crowd-sourced T-shirt designs. In few cases, contests have awarded much larger prizes, such as the $1 million paid in 2009 by Netflix for improving movie recommendations, and the multimillion dollar prizes for space exploration from the X Prize Foundation and NASA's Centennial Challenge. (Like open source, crowd-sourcing is an area where open innovation and user innovation provide distinct but overlapping perspectives and predictions.)

If there is increased interest and research on commercializing internal innovations, the emphasis is most likely less on the R&D-intensive multinationals (like Xerox or IBM) and more on small start-ups trying to solve the Teece (1986) commercialization dilemma. Open innovation practices would be applicable to incubators and programs such as the US Small Business Innovation Research grants, which seek to help fledgling entities unlock value from their innovations, whether by bringing products to market or licensing their technologies or components to other firms.

Levels of Analysis

In principle, open innovation is a form of

economic organization that has implications for industries, sectors, and national economies. However, both as conceived and originally researched, the focus has almost entirely been about dyadic exchange (West, Vanhaverbeke, and Chesbrough, 2006). Often, the emphasis is on the firm and its make versus buy decision to procure external innovations from the market (Laursen and Salter, 2006), or the analogous considerations on offering internal innovations to the market. In a transactional perspective, studies consider the search, evaluation, contracting and enforcement activities of in- or out-licensing technologies (e.g., Chesbrough and Crowther, 2006).

While this exchange normally takes place between two firms, even this dyadic perspective can be expanded in several ways. The supplier of external innovations to a firm may not simply be another company, it may be a nonprofit entity such as a university or government, R&D joint venture, or an individual (Chesbrough, 2003b; West and Gallagher, 2006). Increasingly, both open innovation and user innovation researchers are interested in these alternative suppliers of innovation to firms (West and Lakhani, 2008). These entities (e.g., university technology licensing offices) may have the same profit-maximizing motives as other innovation suppliers, but often the nonprofit actors are driven by nonpecuniary motives such as reputation.

However, the adoption or practice of open innovation can be studied across many other levels of analysis beyond the firm (Table 1). The dyadic perspective has been extended horizontally and vertically to compare and contrast the innovation flows across a business ecosystem. Open innovation may be an integral part of a value network of relationships between firms in a given industry or sector (Vanhaverbeke, 2006). Alternately, horizontal contrasts can be drawn between the choices of similarly situated firms in a given market (Christensen *et al.*, 2005). These two perspectives can be combined to contrast the performance of rival value networks practicing varying levels of vertical integration versus open innovation, but to the author's knowledge this has not yet been done.

Open innovation also has implications for entire industries, sectors, or national economies. Industries that develop a supply of (or demand for) external innovations may de-integrate around interlocking relationships of innovation buyers and innovation merchants. The supply of such innovations may be fueled by nonprofit or government innovation benefactors, such as the Wellcome Trust or the National Renewable Energy Laboratory. Meanwhile, national governments (particularly in the European Union) have taken an interest in policies that promote or discourage open innovation

Empirical Studies

As with any new theoretical framework or paradigm, case studies were prominent in the open innovation research immediately following Chesbrough (2003a); space limitations prevent mentioning more than a handful. Chesbrough himself provide key case studies involving IBM, Intel, Procter & Gamble, and Xerox (Chesbrough (2003a); Chesbrough and Rosenbloom, 2002). West and Gallagher (2006) considered firm open innovation strategies for 10 open source software projects – involving both firm- and individual-level contributions – as well as individuals contributing modifications ("mods") to role-play video games. Christensen *et al.* (2005) contrasted the outside-in, inside-out and vertically integrated innovation strategies of consumer electronics manufacturers in response to the shock of a radical innovation, the Class D digital amplifier.

Another research design commonly used early in a new research paradigm is an exploratory qualitative research. In one such study, Chesbrough and Crowther (2006) identified 40 early adopters of open innovation outside "high tech" (i.e., IT or pharmaceutical industries), and then conducted semistructured interviews with 12. Their study suggests different motivations and strategies used by firms for identifying, procuring and integrating external source innovations. They also found that firms seeking external innovations sought new applications of proven technologies to their industries rather than risking investments in unproven "new to the world" technologies.

Broader generalizability requires large-scaled studies of a representative population of firms or other innovation participations. One of the

Table 1 Levels of analysis in open innovation research.

Level of Analysis	*Definition of Adoption*	*Possible Phenomena*
Individual	Firms source innovations from nonemployees	Crowd-sourcing Video game "mods"
Community	Firms leverage a process of external innovation by communities	Open source software Open science
Organization	Firms obtain external innovations from other organizations	Component business models Patent licensing Sponsored university research
Value network	Firms cooperate across a value network to create and commercialize innovations	System integration Ecosystem management
Industry/sector	Open innovation is commonplace by firms across an entire industry or sector	Personal computers Biotechnology
National institutions	Government agencies adopt policies supporting open innovation	Patent reform National laboratories

Source: Updated from West *et al*. (2006) and West and Lakhani (2008).

first such tests of open innovation came with Laursen and Salter (2006), who used responses from 2707 firms in the UK Innovation Survey to establish a curvilinear relationship between the level of innovation search and innovation performance. Subsequent researchers used national innovation surveys in other European countries including Belgium, Netherlands, Germany, and Italy. Such surveys provide systematic evidence of open innovation across a broad range of national contexts, but the nature of the existing survey instruments limits the inferences regarding causal relationships.

Future Directions

By its nature, the research identifying open innovation both encourages broader use of open innovation and changes how it is practiced. In particular, new approaches and mechanisms for open innovation are providing new phenomena for study.

However, there remains a wide range of industries and contexts to be studied. This includes low-tech industries, services, and small and medium enterprises. It also includes studies of firms and policies in other national contexts, beyond the existing settings of Europe and North America. Researchers will also find fertile opportunities for research in companies and industries that are transforming from vertically integrated to open innovation (or balanced innovation) approaches to produce new products and services.

Most research on open innovation has focused on where it occurs, not where it does not. To identify the preconditions of open innovation, researchers might consider predictions from transaction cost economics. Has the Internet reduced search costs so much that open innovation is appropriate in all national contexts? Or does open innovation occur only in those countries where there are properly functioning IP institutions, and with a pool of small innovative companies that feel they are able to bargain fairly with larger firms that have the ability to commercialize?

Longer term, an important area of future research is developing a detailed understanding of how firms internally organize to develop and commercialize using open innovation, both building upon and updating what we know about organizing for innovation under existing models of industrial innovation. We know a lot about cross-functional teams, lightweight and heavyweight project managers, stage-gate and waterfall scheduling models, but we lack a similar understanding of best practice within firms practicing open innovation.

Researchers also have the opportunity to investigate the overlap of open innovation and other theories. For example, there are numerous opportunities to research phenomena (such as crowd-sourcing) at the intersection of user innovation and open innovation. The two streams of research have different assumptions, predictions, and normative and policy recommendations, but appear to offer congruent suggestions for enabling and identifying sources of external innovations.

As with any theory, it is important to define the boundary conditions of open innovation, either as a phenomenon, or as a "new paradigm" (cf. Chesbrough, 2006) for explaining existing phenomena, particularly regarding the role of small companies as suppliers of innovation. While Chesbrough (2003a) emphasized open innovation as a new paradigm for large firm innovation, does open innovation have something to say to small companies, who were unlikely to fully integrate even before Chesbrough's research? Similarly, if a large company (like Cisco or Google) scans for external sources of innovation – but commercializes that by buying its creator – what contribution does the open innovation perspective provide? Is that contribution any greater if (as Chesbrough and Crowther, 2006 suggest) firms buy exclusive rights to a small firm's technology? And what are the implications of such open innovation strategies – exclusive and nonexclusive – for the creation, entry, and growth of innovative start-up companies?

Acknowledgements

My appreciation to Fern Evitt of the University of Auckland for sharing the open innovation literature review from her doctoral dissertation.

Bibliography

Chesbrough, H. and Rosenbloom, R.S. (2002) The role of the business model in capturing value from innovation: evidence from Xerox corporation's technology spin-off companies. *Industrial and Corporate Change*, **11** (3), 529–555.

Chesbrough, H. (2003a) *Open Innovation: The New Imperative for Creating and Profiting from Technology*, Harvard Business School Press, Boston.

Chesbrough, H. (2003b) The era of open innovation. *Sloan Management Review*, **44** (3), 35–41.

Chesbrough, H. (2006) Open innovation: a new paradigm for understanding industrial innovation, in *Open Innovation: Researching a New Paradigm* (eds H. Chesbrough, W. Vanhaverbeke and J. West), Oxford University Press, Oxford, pp. 1–12.

Chesbrough, H. and Crowther, A.K. (2006) Beyond high tech: early adopters of open innovation in other industries. *R&D Management*, **36** (3), 229–236.

Christensen, J., Frøslev, M., Holm, O. and Kjær, J.S. (2005) The industrial dynamics of open innovation — evidence from the transformation of consumer electronics. *Research Policy*, **34** (10), 1533–1549.

Enkel, E., Gassmann, O. and Chesbrough, H. (2009) Open R&D and open innovation: exploring the phenomenon. *R&D Management*, **39** (4), 311–316.

Laursen, K. and Salter, A. (2006) Open for innovation: the role of openness in explaining innovation performance among UK manufacturing firms. *Strategic Management Journal*, **27** (2), 131–150.

Teece, D. (1986) Profiting from technological innovation: implications for integration, collaboration, licensing and public policy. *Research Policy*, **15** (6), 285–305.

Vanhaverbeke, W. (2006) The inter-organizational context of open innovation, in *Open Innovation: Researching a New Paradigm* (eds H. Chesbrough, W. Vanhaverbeke and J. West), Oxford University Press, Oxford, pp. 205–219.

von Hippel, E. (2005) *Democratizing Innovation*, MIT Press, Cambridge, MA.

West, J. (2006) Does appropriability enable or retard open innovation?, in *Open Innovation: Researching a New Paradigm* (eds H. Chesbrough, W. Vanhaverbeke and J. West), Oxford University Press, Oxford, pp. 109–133.

West, J. (2009) Policy challenges of open, cumulative and user innovation. *Washington University Journal of Law and Policy*, **19**, 17–41.

West, J. and Gallagher, S. (2006) Challenges of open innovation: the paradox of firm investment in open-source software. *R&D Management*, **36** (3), 319–331.

West, J. and Lakhani, K. (2008) Getting clear about the role of communities in open innovation. *Industry and Innovation*, **15** (2), 223–231.

West, J., Vanhaverbeke, W. and Chesbrough, H. (2006) Open innovation: a research agenda, in *Open Innovation: Researching a New Paradigm* (eds H. Chesbrough, W. Vanhaverbeke and J. West), Oxford University Press, Oxford, pp. 285–307.

bass model

Christophe Van den Bulte

Model Structure

The Bass model is a mathematical model of new-product diffusion (*see also* DIFFUSION OF INNOVATION). It represents how the numbers of people (or other purchasing units, such as households or firms) that have adopted a new product or technology increase over time. The model represents people as being in one of two possible two states at any time: either they have adopted, or they have not adopted. Once they have adopted, they cannot disadopt.

The model can be represented as a differential equation showing what drives the number of adopters to change over time, or as the solution to that equation showing how the cumulative number of adopters increases over time. In this article, the cumulative number of adopters at time t is denoted as $N(t)$ and its first-order differential as $dN(t)/dt$. Note, the latter captures the number of adoptions taking place in the very short time interval $[t, t + dt]$ divided by the length of that interval. For simplicity, $dN(t)/dt$ is often referred to as simply the *number of adoptions taking place at time t*.

The model has three parameters, typically identified as p, q, and m. Parameter p is typically called the *parameter of innovation* or *external influence*. Parameter q is typically called the *parameter of imitation* or *internal influence*. Parameter m is typically called the *market potential*, *market size*, or *diffusion ceiling*. The reason for these names stems from the behavioral rationale for the model, which is conveyed through its differential equation form.

Differential equation. The differential equation form of the model represents the number of adoptions at time t as follows:

$$\frac{dN(t)}{dt} = \left[p + \frac{qN(t)}{m}\right][m - N(t)] \qquad (1)$$

The second term in brackets on the right-hand side represents how many people still have to adopt at time t, by subtracting the cumulative number of adopters $N(t)$ from the number of people who will ultimately adopt, m. The first term in brackets represents the rate at which people transition from the nonadopter to the adopter state, that is, the rate at which the number of adopters increases.

The transition rate has two components. The first is p, capturing all time-invariant elements that may induce people to or inhibit them from adopting quickly. These may include the product's price, the level of marketing communication support it receives, the extent to which it is easily available in stores, product characteristics such as superior functionality and low risk profile, as well as many economic, demographic, and cultural characteristics of the population of ultimate adopters. The second component of the transition rate is $qN(t)/m$, that is, the parameter q multiplied by the fraction of the market that has already adopted. This second component captures how the rate at which people adopt is affected by how many have already adopted in the past, which may stem from imitation, word of mouth, social contagion, network effects, or other endogenous feedback mechanisms. This interpretation is why q is often referred to as the *parameter of imitation*.

Some scholars interpreting p as capturing only the effect of mass media and marketing communications refer to p as the *coefficient of external influence* (capturing the effect of outside influences acting on the group of consumers) and to q as *the coefficient of internal influence* (capturing social influence effects operating within the group of consumers). However, research documenting that p varies as a function of adopter characteristics (e.g., national culture) suggests that these interpretations and labels do not capture the entire range of factors captured by p.

Diffusion as a function of time. Whereas the differential equation clearly conveys the behavioral rationale of the model, it does not provide a convenient expression directly showing how the number of adopters increases over time. The latter can be obtained by solving the equation as a function of time only. Assuming that $t = 0$ is the time of launch such that $N(0) = 0$, the solution is

$$N(t) = m\frac{[1 - e^{-(p+q)t}]}{[1 + (q/p)e^{-(p+q)t}]} \qquad (2)$$

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Mathematical analysis of this equation provides additional insight into how the model parameters affect the diffusion curve, with time on the horizontal axis and the cumulative number on the vertical axis. First, m simply elongates or shrinks the entire curve vertically, without affecting its shape, and so "scales" the diffusion ceiling or market size. Second, the sum of p and q elongates or shrinks the entire curve horizontally, without affecting its shape, and so "scales time." Finally, the ratio of q and p determines the shape of the curve. Whenever q is larger than p, the curve will have an S-shape, and the larger the (q/p) ratio is, the more pronounced the S-shape will be. The time of inflection when the S-curve goes from being convex to being concave, which corresponds to the time at which the largest number of people are adopting, is given by $t^* = [\ln(q/p)]/[p + q] > 0$ when $q > p$. When q is equal to or smaller than p, in contrast, the curve of cumulative adoption is concave without any inflection point, and the number of adoptions taking place at any point in time is the highest at $t = 0$ and then decreases over time.

Figures 1 and 2 illustrate how q/p drive the shape of the curve. Figure 1 shows that, when q is larger than p, the cumulative number of adopters $N(t)$ follows the type of S-curve often observed for really new product categories (*see* TECHNOLOGY S-CURVE). When q is smaller than p, the cumulative number of adopters follows an inverse J-curve often observed for less risky innovations. Note, the curves in Figure 1 are rescaled to be cumulative penetration curves, that is, they show $N(t)$ divided by m. Figure 2 shows the corresponding patterns for $dN(t)/dt$, again divided by m, which can be interpreted as the proportion of new adoptions occurring at time t. Note that an S-curve for the cumulative proportion of adopters (Figure 1) corresponds to the familiar bell-shaped curve for the noncumulative proportion of adopters (Figure 2).

Typical values for parameters p and q. The model has been applied mostly to consumer durables, high-technology products, and pharmaceuticals. For such products exhibiting considerable price or risk barriers, the typical p is about 0.01–0.03 and the typical q is about 0.30–0.40, both measured on an annual scale (Sultan, Farley, and Lehmann, 1990; Van den Bulte and Stremersch, 2004). That q is markedly higher than p is consistent with the distinct S-shape in the diffusion curve of many such products. For consumer packaged goods and other low-risk items such as music CDs or movies with well-known performers, p is often higher (e.g., 0.10 on a monthly scale) and q often is zero. Note, the averages just reported are based mostly on data from the United States, and there is evidence that in emerging economies, values of p tend to be lower and those of q tend to be higher than those just mentioned (Talukdar, Sudhir, and Ainslie, 2002; Van den Bulte and Stremersch, 2004).

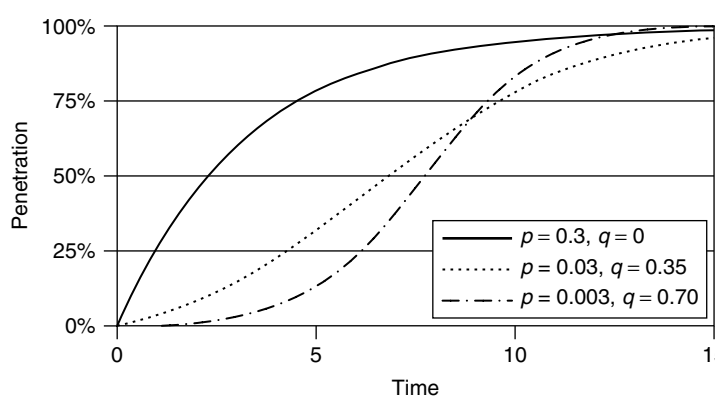


Figure 1 The Bass model fits the classic S-shaped cumulative diffusion curves.

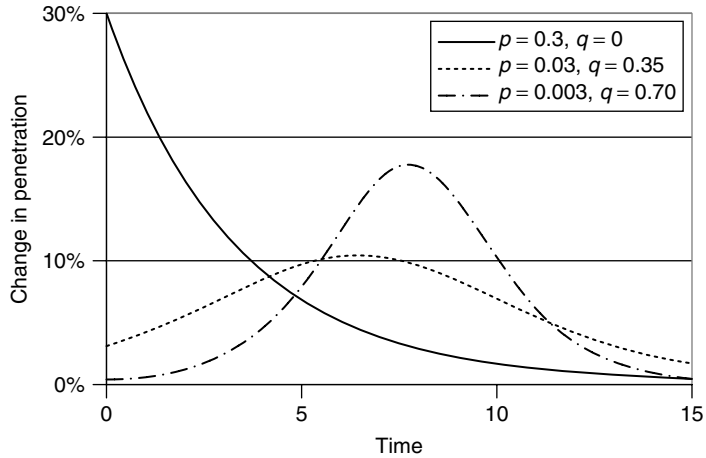


Figure 2 The Bass model fits the classic bell-shaped diffusion curves.

Interpretation

Meaning of the parameters. The three parameters not only capture the overall shape of the diffusion process but also have an appealing behavioral interpretation: m represents the market size, q captures word of mouth and other types of social contagion, and p captures all time-invariant influences associated with speedy diffusion. Since the two rate parameters p and q allow two distinct types of influence to operate, the Bass model is also referred to as the *mixed-influence model of diffusion* (Mahajan and Peterson, 1985).

Note that, with this interpretation, the model makes several assumptions. First, the market size m does not change over time. Second, any marketing-mix influences that vary over time are ignored. Third, all people who have already adopted are allowed to influence all people who have yet to adopt, which implies either that everyone interacts with everyone else or that contacts in the population are totally random. Fourth, repeat purchases, multiunit purchases, and disadoptions all fall outside the purview of the model. These restrictive assumptions can be relaxed, but at the cost of greater complexity (Mahajan, Muller, and Bass, 1993).

There is disagreement, not to say confusion, about a fifth assumption: does the model assume a homogeneous population or does it allow for different types of customers to exist? This issue is of fundamental importance because it relates to the presence or absence of customer segments, and is discussed in the next two sections.

Innovation versus imitation. It is tempting to think about the model as representing a diffusion process not only with two distinct influences at work but also with two distinct customer groups or segments: "innovators" unaffected by social influence and "imitators" affected by social influence. Considerable confusion has arisen on this issue.

The model derivation presented above recognizes no differences among potential adopters; that is, though the model recognizes that different people will adopt at different times, the derivation assumes a homogenous market where everyone is equally exposed and sensitive to both independent and social influences. It is clear from Equation (1) that the only people adopting without being exposed to any social contagion are those adopting at $t = 0$. The fraction of people adopting at that time is $p \cdot dt$, which is extremely small since the model is technically a continuous-time model where dt is extremely small. So, even though p can validly be interpreted as a coefficient of innovation, it is not the fraction of people adopting independently. The latter is infinitesimally small according to the model's mathematical structure. If one starts with the assumption that two distinct groups of customers exist, one

adopting with transition rate p and the other with transition rate $qN(t)/m$, one obtains a different mathematical model (Van den Bulte and Joshi, 2007).

So, one cannot interpret the Bass model as a two-segment model. However, it is possible to interpret it as depicting a situation where everyone can, in principle, adopt either independently or through contagion but, at the time of adoption, some people adopt only because of independent forces captured by p and others adopt only because of social influence captured by $qN(t)/m$ (Daley, 1967; Taga and Isii, 1959). This alternative scenario, where no ex ante segments exist but where some adoptions happen independently and others happen only because of social influence—such that adoptions can be said to be either independent or imitative ex post (once they have occurred)—actually is consistent with the Bass model. Under this interpretation of the model, the fraction of independent adoptions taking place at time t decreases continually over time (Mahajan, Muller, and Srivastava, 1990). The fraction of all independent adoptions from the time of launch until full diffusion is achieved, say ϕ, equals

$$\phi = \frac{p}{q} \ln\left[1 + \frac{q}{p}\right] \tag{3}$$

Note, this fraction is a function only of the q/p ratio. For independent adoptions to account for only 2.5% of all adoptions, as some people have proposed to make the definition of innovators as people adopting independently, consistent with Rogers' classic definition of innovators as the first 2.5% of people who adopt (Rogers, 2003), the q/p ratio should equal 215. The latter value is well above the typical value (Van den Bulte and Stremersch, 2004), further underscoring the fact that the concepts of "innovators" and "innovative adoptions" have very specific meanings within the Bass model structure that need not correspond to people's intuitions.

Imitation versus heterogeneity. The discussion so far has followed traditional lines, and presented the model as the mathematical representation of a process among homogenous customers exposed to independent and imitative influences. Alternative interpretations of the mathematical formulas are possible, however. Like many other mathematical models, the Bass model is overdetermined. In other words, different sets of assumptions can lead to exactly the same mathematical expression. Today, researchers recognize that Equation (2) corresponds not only to a diffusion process in a homogenous population driven by time-invariant independent forces and imitation but also to a diffusion process in a heterogeneous population driven by time-varying independent forces without imitation.

One alternative derivation was presented by Bemmaor (1994). He showed that if people adopt according to a shifted Gompertz process and their tendency to adopt early is exponentially distributed, the resulting expression for the expected number of cumulative adopters is identical to Equation (2). Another alternative is based on the logic of adoption thresholds, well established in economics. The most obvious scenario is one where the price for the new product decreases over time (e.g., because of experience curve effects or increasing competition), and people adopt as soon as the price falls below their maximum willingness to pay or reservation price. If prices come down exponentially (which they tend to do) and the cumulative distribution function of the natural logarithm of reservation prices is $1 - N(t)/m$ where $N(t)$ is defined as in Equation (2), but with the log of price at time t substituted for t, then the cumulative adoption curve will be identical to Equation (2). Assuming such a distribution of threshold prices is consistent with the distribution of income in many countries.

Because the Bass model is overdetermined, one cannot be very confident about what it means when the model fits data well. It might be that the standard mixed-influence process is at work, but one cannot make that conclusion confidently unless one has ruled out the other interpretations. For instance, a study investigating how the q/p ratio in particular countries varied systematically with their income distribution and with several dimensions of national culture associated with social influence found support for both the mixed-influence and the income-heterogeneity interpretations of the model (Van den Bulte and Stremersch, 2004). This finding, together with other studies showing that the evidence of

social contagion may vanish once one controls for changes in the marketing mix, suggests that one must be cautious in drawing conclusions on what drives the diffusion process. While it is true that S-shaped diffusion curves and high values of the q parameter in the Bass model are consistent with imitative processes, they may also result from systematic changes in prices and other factors.

Estimation

Because Equations (1) and (2) are not linear in the parameters, one cannot get direct estimates of p, q, or m using ordinary least squares (OLS) regression. However, one can get direct estimates using maximum likelihood estimation or nonlinear least squares regression. The latter is the most popular estimation today. It is quite easy to implement with standard statistical software and can even be done using Excel Solver. More sophisticated techniques exist as well (Mahajan, Muller, and Wind, 2000).

Obviously, a problem occurs when one wants to use the model for prelaunch forecasting and quantify the model parameters before any sales or adoption data is available. To circumvent this problem, one must use other sources of information, such as managerial judgments, inputs from customer surveys, or the parameter estimates from similar products and countries for which more data is available.

Problems are very likely to occur even after launch, especially very early in the diffusion process, when only limited sales or adoption data is available. In some cases, it is simply impossible to get reasonable statistical estimates. In other cases, the estimates one obtains are reasonable but vary either randomly or systematically as one adds more data points to the data set. If one does not have the luxury to wait for more data to come in, one can again resort to managerial judgments, inputs from customer surveys, or parameter estimates from similar products and countries. In addition, one can also try to combine multiple short data series from multiple products or countries and "borrow strength" across them using random effects models (either Bayesian or non-Bayesian).

So, although nonlinear regression is the most commonly used method to estimate the Bass model, there are situations where other approaches may be required. Estimation issues are discussed at greater length in Mahajan, Muller, and Bass (1993) and Mahajan, Muller, and Wind (2000).

Uses

The model can be used to predict future adoptions or first-time purchases. It can also be used to describe and explain patterns of diffusion. When extended with marketing-decision variables such as price or advertising, the model can also be used to provide advice on how the marketing mix should change over time to maximize profits.

Forecasting. Sales forecasting is one of the model's most popular uses (*see* NEW-PRODUCT FORECASTING). Both experience and research indicate that the model can provide reasonably good short-term forecasts, but that one can typically do better by combining several forecasts obtained through different models or forecasting techniques (Meade and Islam, 1998).

The model also allows one to forecast what the maximum amount of first-time purchases will be and when they will occur. This information can be valuable for planning production capacity. As already mentioned, the time at which the largest number of people are adopting is given by $t^* = [\ln (q/p)]/[p + q]$ when $q > p$. The maximum number of adoptions taking place throughout the entire diffusion process is then $dN(t)/dt$ evaluated at t^*, which equals $[m(p + q)^2]/(4q)$.

Managers and analysts interested in applying the model to predict future sales must be aware of some issues. First, the model is meant to forecast only adoption or initial purchases, not repeat purchases. Second, the simplicity of the model comes at the price of some strong assumptions. Perhaps the most important one for forecasting purposes is that the potential market size is constant over time, and hence not a function of changes in marketing strategy or in macroeconomic conditions. A third issue to be aware of is that the Bass model can accommodate various smooth diffusion trajectories, such as S-shaped diffusion curves consistent with bell-shaped adoption curves, but not the phenomenon of market chasms between

early and late market segments corresponding to a double-humped adoption curve. Of course, the model can be extended to relax those assumptions, but that will typically make the model more complex and hence less practical. So, depending on whether repeat purchases are important within the forecasting window, whether one can assume the potential market size to remain fairly constant, and whether one believes market chasms to be very likely, the Bass model may or may not be quite appropriate.

Describing and explaining patterns of diffusion. Managers and academics sometimes use the model as a means to quantify, through numerical values of the model parameters, the pattern of diffusion for multiple products, multiple industries, or multiple countries. When comparing diffusion paths for several products or markets, looking at raw data series need not always be informative and patterns may become clearer after plotting parameter values in a two-dimensional map or displaying the smooth best-fit curves rather than the jagged raw data series (of course, color coding can further enhance the intelligibility of one's graphs).

The model can also be quite useful when one is interested in explaining variations across products and markets. One can obtain estimates for p, q, and m (or better, m divided by the size of the population of individuals, households, or industrial establishments), and then use standard statistical techniques to see if they are systematically related to market and product characteristics. Elements that have been shown to affect the diffusion path include average income, income heterogeneity, media and distribution infrastructure, legislation and regulation supporting or inhibiting the uptake of the product, and the presence of competing standards (Talukdar, Sudhir, and Ainslie, 2002; Van den Bulte and Stremersch, 2004).

Of course, companies and consultants may want to build industry-specific databases and look for patterns across products most relevant to their own line of business. In the pharmaceutical industry, for example, researchers have investigated to what extent the coefficients of innovation and imitation are related to drug efficacy, the presence of side effects, being aimed at treating a life-threatening medical condition, being aimed at treating an acute versus chronic medical condition, lack of competition, and the amount of marketing support at launch. Such industry-specific analyses can provide valuable insights to marketing decision makers.

The model has been extended in many ways to account for more complex phenomena than the original model can account for, such as growth in the market potential, influences across competing or complementary products, the presence of supply constraints, and replacement purchases (Mahajan, Muller, and Wind 2000). Two kinds of extensions, accounting for cross-country influences and for marketing-mix influences, can have particularly important implications for how to market new products.

Improving marketing strategies. Comparisons of diffusion processes across industries or countries provide insights not only about the likely drivers of diffusion but also about what type of industry or country is likely to experience a slow or fast diffusion. The latter can be of immediate use to companies and guide decisions on which markets to enter early versus late, or on how much to spend on marketing in each target market (*see also* LAUNCH STRATEGIES).

Some studies have extended the model to allow the penetration rate or installed base in one country to influence the subsequent diffusion rate not only in the same country but also in other countries (for a review, see Mahajan, Muller, and Wind, 2000). The evidence suggests that such cross-country influences exist, and that they are asymmetric. This, in turn, has been shown to be an important consideration for how multinational companies should launch their new products. Keeping everything else constant, the presence of influential countries supports a "waterfall strategy," where firms enter markets sequentially from leaders to followers rather than a "sprinkler strategy," where firms enter markets simultaneously, because capitalizing on the free cross-country influence from leading markets allows firms to penetrate follower markets more rapidly or with less marketing support. The logic is exactly the same as in viral strategies where firms target opinion leaders, in the hope that their influence will then snowball through the market and convert other customers at no cost to the firm.

The Bass model has also been extended with marketing decision variables, such as price and advertising. Unsurprisingly, empirical research suggests that marketing-mix variables affect the diffusion path. There is little to no consensus, however, on the specific mathematical form in which such decision variables should enter the model: through p only, through p and q equally, or through m (Mahajan, Muller, and Bass, 1993; Mahajan, Muller, and Wind, 2000). Some researchers feel that different marketing variables should enter the model in different ways, for example, advertising through p (which makes sense if one believes that advertising affects how quickly people decide to adopt, without affecting how many people will ultimately adopt) and price through m (which makes sense if one believes that price affects the fraction of the population that can afford the product but does not affect how quickly one adopts given that one will adopt ultimately). Unfortunately, empirical research provides no clear indication that one particular model specification is better than any other. This lack of guidance from empirical research is particularly problematic as different model specifications, with the marketing-mix variable entering one way rather than another, can have very different implications for how firms should change their marketing mix over time. For instance, models in which advertising boosts p tend to favor strategies where one decreases advertising over time. This is intuitively appealing: one first advertises heavily to get the first adopters, and can then cut advertising expenditures as word of mouth becomes a major influence in converting nonadopters into adopters. In contrast, models in which advertising acts as a multiplier on the entire diffusion rate $[p + qN(t)/m]$, which is equivalent to amplifying p and q to the same extent, tend to favor strategies where one first increases advertising over time and then decreases it. This too makes sense, as it implies that one advertises the most when the most adoptions take place. So, we have two very different strategies that are both logically defensible, but on different grounds. The sensitivity of model-based recommendations to model assumptions, combined with the lack of empirical evidence on which set of model assumptions is empirically most defensible, severely limits the practical usefulness of such theoretical work to date.

Origins and Naming Conventions

The model is named after Frank Bass, who introduced the model to the marketing community in an article published in 1969. The model, however, had been presented by several others before, with Taga and Isii (1959) arguably being the first, and with the most famous sources being Mansfield (1961) in economics and Coleman (1964) in sociology. The article by Bass (1969), however, was the first to actually fit the model to data and estimate values for p, q, and m. It also showed that the model could be used for forecasting purposes, rather than for only descriptive purposes. Today, most marketers refer to the model simply as the Bass model. Some sources also refer to it as the mixed-influence model of diffusion (Mahajan and Peterson, 1985).

Acknowledgments

The text benefited from helpful suggestions by Professor Renana Peres, Hebrew University of Jerusalem.

Bibliography

Bass, F.M. (1969) A new product growth model for consumer durables. *Management Science*, **15**, 215–227.

Bemmaor, A.C. (1994) Modeling the diffusion of new durable goods: Word-of-mouth effect versus consumer heterogeneity, in *Research Traditions in Marketing* (eds G. Laurent, G.L. Lilien, and B. Pras), Kluwer Academic Publishers, Boston. pp. 201–223.

Coleman, J.S. (1964) *Introduction to Mathematical Sociology*, The Free Press of Glencoe, London.

Daley, D.J. (1967) Concerning the spread of news in a population of individuals who never forget. *Bulletin of Mathematical Biophysics*, **29**, 373–376.

Mahajan, V., Muller, E., and Bass, F.M. (1993) New-product diffusion models, in *Marketing* (eds J. Eliashberg and G.L. Lilien), North-Holland, Amsterdam, pp. 349–408.

Mahajan, V., Muller, E., and Srivastava, R.K. (1990) Determination of adopter categories by using innovation diffusion models. *Journal of Marketing Research*, **27**, 37–50.

Mahajan, V., Muller, E., and Wind, Y. (eds) (2000) *New-product Diffusion Models*, Kluwer, Boston.

Mahajan, V. and Peterson, R.A. (1985) *Models for Innovation Adoption*, Sage Publications, Newbury Park.

Mansfield, E. (1961) Technical change and the rate of imitation. *Econometrica*, **29**, 741–766.

Meade, N. and Islam, T. (1998) Technological forecasting—model selection, model stability, and combining models. *Management Science*, **44**, 1115–1130.

Rogers, E.M. (2003) *Diffusion of Innovations*, 5th edn, Free Press, New York.

Sultan, F., Farley, J.U., and Lehmann, D.R. (1990) A meta-analysis of diffusion models. *Journal of Marketing Research*, **27**, 70–77.

Taga, Y. and Isii, K. (1959) On a stochastic model concerning the pattern of communication: diffusion of news in a social group. *Annals of the Institute of Statistical Mathematics*, **11**, 25–43.

Talukdar, D., Sudhir, K., and Ainslie, A. (2002) Investigating new product diffusion across products and countries. *Marketing Science*, **21**, 97–114.

Van den Bulte, C. and Joshi, Y.V. (2007) New product diffusion with influentials and imitators. *Marketing Science*, **26**, 400–421.

Van den Bulte, C. and Stremersch, S. (2004) Social contagion and income heterogeneity in new product diffusion: a meta-analytic test. *Marketing Science*, **23**, 530–544.

brainstorming

Jacob Goldenberg and David Mazursky

Brainstorming—A Critical Review

The knowledge underlying the concept of "brainstorming," the relative ease of its operation and assimilation in organizational contexts, and the social benefits incurred in brainstorming encounters have created a ubiquity of brainstorming group sessions for problem-solving discussions in large organizations (*see also* CREATIVITY; IDEA MANAGEMENT). Advertising agency staff meet for brainstorming sessions to develop creative concepts or new advertising strategies. Engineers meet to find a solution to problems that "arrest R&D progress" and even CEOs initiate encounters with managers from various levels to review and identify new ideas for the advancement of the organization and its components. These sessions are either planned or improvised, and although sometimes guided by a professional facilitator or by one of the group participants, quite often, the sessions are unguided. The high degree of variety in the nature and management of these discussions may have even contributed to the high acclaim received by the genre of brainstorming sessions in popularity polls among methods of group thinking.

The origin of brainstorming is hardly contested. At the Disney Studios, brainstorming was an accepted method to inspire professional creativity by encouraging interaction and teamwork. Osborn (1957) conceptualized the approach, setting down guidelines to create mental storming. There are a number of assumptions underlying Osborne's approach:

1. *People are naturally creative*: Unfortunately, the ties that bind us to our routines, and the demanding pace of life, inhibit us from flourishing creatively and generating innovative ideas. Confronted with analogies, we are released from our bonds—exposure to an analogy "disrupts" our routine reasoning and opens our mind to associations that lead to the production of original thought.
2. *Synergies*: A group of people thinking together is superior to a single person thinking on his or her own. Osborne advocated that "individuals operating in a brainstorming group suggest twice as many ideas as individuals working on their own."
3. *Deferred judgment*: If we eliminate the requirement to pass immediate judgment as ideas are spoken, we can gradually accumulate a pool of high-quality and original ideas which are subsequently filtered. Related to this assumption is the adage "no line of inquiry should be ruled out".
4. *Quantity leads to quality*: The more we increase the number of ideas, the greater the probability of achieving a more qualitative set of ideas after filtering. As Nobel Prize winner Jonas Pulling said, "The best way to get a good idea is to get a lot of ideas".

These assumptions (although formalized after the method came into practice) are the foundation on which the brainstorming method evolved in two major stages:

1. *Conceptual brainstorming*: In this first stage, a group of individuals advance ideas in no particular order, and no criteria for judgment of their merits is applied. Every idea is deemed good, and the more diverse the ideas the better. The expression of even wild and seemingly illogical ideas is invited. When participants listen to the ideas of their fellow group members, they are stimulated to create new ideas in new directions. Osborne advocated that fragments of ideas or thoughts are also welcome, as they may produce a good idea in the mind of another participant. Beyond managing the discussion, the aim of the facilitator, in this stage, is to create a pleasant atmosphere of deferred criticism and to encourage diversity of thought. As Doug Floyd noted, "You don't get harmony when everybody sings the same note". Concurrently with bouncing ideas to each other, group members also respond to the ideas and suggestions of their fellow members, thus refining ideas, building new perceptions based on other ideas, or connecting a number of ideas together to consolidate a new perception.
2. *Screening*: In this stage, tens (or sometimes hundreds) of ideas are filtered to produce

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

a reduced set, which is subsequently examined and tested for economic feasibility and value. Underlying this stage are judgment and criticism, of the type formerly deferred in the first stage. According to the classical approach to brainstorming, screening should be performed by persons other than group members involved in the generation of the ideas. This ensures that the members of the filtering team are not captives of the perceptions or experiences of the first, conceptual stage. The importance of criticism at this stage is recognized, and its contribution is considered more significant after ideas have been generated during the first stage.

Brainstorming has captured the hearts of organizations, industries, and business concerns, many of which proudly announce their adoption of this practice, although, in practice, strict application of its guidelines has not been universal. Once brainstorming became a conventional practice and it seemed that nothing could be added to the original formula, the academic world began to examine the essential quality of the phenomenon. In the 1980s and 1990s, the method's effectiveness and relevant success factors were the subject of an increasing number of research studies. Examples of questions highlighted in this body of research are the search for the optimal number of group members and the optimal duration of the brainstorming session. The central question was, "What is brainstorming's real contribution compared to results of a nominal group (a group of individuals who think alone, with no contact among members)?"

One of the prominent findings was the absence of a clear advantage of the brainstorming group relative to the results of individuals who worked alone. This finding appeared repeatedly until no doubt was left in the minds of investigators: a brainstorming session does not generate more ideas or greater creativity in comparison to groups of individuals working with no contact among them (Diehl and Stroebe, 1987, 1991). Is this just another case of detachment of academic research producing surprisingly precise, though irrelevant, results from the reality in the field? If this were a case of a big promise that turned into a disappointing fad, we have to face the intriguing fact that this so-called management fad has struck roots and thrived for almost 30 years. Admittedly, even the most complex and difficult management hits do not typically maintain a leading position in popularity polls for over two decades. Something is obviously going on here. The question is, who is wrong?

Research indicates that most brainstorming groups did not generate more ideas than their control groups in which individuals worked alone with no contact between them. Already appearing in 1958 (the year of Osborne's publication), the first study on this topic, by Tyler, Berry, and Block (Diehl and Stroebe, 1987, 1991) empirically proved that solitary subjects produced almost twice the number of ideas as subjects working in groups. Most of the numerous studies carried out over the years unanimously supported this conclusion and replicated results that stood in direct contrast to Osborne's claim. Groups were shown to be detrimental to individual productivity (Sutton and Hargadon, 1996).

1. The quality of the ideas themselves and their originality were also inferior to the ideas generated by individuals working without any group effect.
2. The optimal number of group members for a brainstorming session was found to be three or four, inconsistent with the conventional perception of a larger group.

These findings, which reveal that group processes have an adverse affect on creativity rather than contributing to their enrichment, created serious doubts regarding brainstorming as an effective process. Explanations suggested by researchers for the lack of success of brainstorming in laboratory experiments range from targeting elements of individual behavior in groups to the impact of brainstorming on individual problem-solving processes (Paulus *et al.*, 1993). The following is a list of some of relevant factors that have been proposed:

- *Production blocking*: In the course of idea generation, one person speaks while the others listen. As, at any single moment in time, only a single person can contribute ideas to the group, the scope of the potential contribution of individual group members

is limited. Moreover, listening to others express their ideas makes it harder for individuals to concentrate and develop ideas of their own. Even if they succeed in concentrating on their own ideas, recall ability is diminished.

- *Free riding*: As is often the case in groups in which individual efforts are combined, brainstorming groups are not immune to free-riding attempts. In a group situation, individuals contribute their ideas to a group pool, and, consequently, are granted recognition on a group level. An opportunity is created for some members to hang onto the coattails of other, more productive members, and bask in the recognition won by group efforts without contributing personally. These free riders, who may possibly function as creative individuals in a different setting, either repeat ideas already expressed or avoid participating in the discussion.
- *Distractions*: The flow of ideas spoken aloud overwhelms individuals straining to concentrate and develop their own innovative thought. Repeated interruptions compel them to withdraw into simpler ideas that are better able to withstand the "external noise." Thus individuals are diverted to routine thoughts, mostly in an unconscious process. A related factor affecting the generation of ideas adversely in a group situation is the fact that each individual shares "thinking time" between his or her own thoughts and all the thoughts of his or her fellow members expressed aloud. Thus, group production declines relative to the sum of individual contributions in a nominal group.
- *Deferred judgments create a chaotic world*: In a world with no judgment or criteria for assessment, individuals have no way of knowing if they are "on the right path." Rather than promoting uninhibited thought, the absence of criteria for successful ideas blocks the flow of thought. This creates two phenomena. The first is a type of helplessness and lack of direction that is typical of a soldier who loses his or her bearing on a navigational task. The second is related to the cognitive loss of bearing, whereby random attempts to generate ideas unfounded on prior reasoning typically produce routine, well-rehearsed thoughts. These familiar ideas are more easily produced than elusive, less-developed notions.
- *Fear of assessment*: Apparently, despite instructions to the participants prior to a brainstorming session, a degree of apprehension of negative social feedback and criticism persists, inhibiting members from expressing all their original ideas.

This last proposed explanation of relatively inferior results by brainstorming seems counterintuitive at first glance. From our experience with brainstorming, the ideas generated are so wild and preposterous that they could not be reflecting a fear of criticism. A more careful examination has shown that the reality is more complex: it seems that participants have no consternation about expressing a wild idea when there is no chance of its practicability. The rules of the brainstorming game allow far-fetched ideas, and the worst that can happen is that an idea will be received as a joke and not taken seriously as an idea for implementation. But, in the case of an original idea that is genuinely directed at the solution of the problem, criticism may be expressed (and most probably, not forgotten), as it threatens the consensus built on the solutions proposed up to that point. This creates a mechanism that filters out the most practical and dauntless of the original ideas, encouraging more banal ideas on the one hand, and wilder, impractical ideas on the other.

Does all the above lead to the conclusion that brainstorming is like a storm in a teacup? This is not necessarily so. Despite the laboratory experiments that invalidate the effectiveness of brainstorming, field results indicate widespread adoption of the method. If indeed such a large number of people cannot be wrong, then the widespread adoption and persistence of brainstorming as an organizational practice requires an attempt to reconcile this inconsistency. Some of the explanations for the contradiction between research findings and practice may lie in what is known as the "*illusion of group effectiveness*."

1. *Lack of distinction between process and outcomes*: The participants' ability to disti-

nguish between their satisfaction in the process and its outcomes is distorted by the magic spell of their experience. In other words, the participants' reported satisfaction in the process is derived from their participation in the brainstorming process, leading them to overevaluate the results compared with individuals who worked on their own and generated the same ideas. Participants in a brainstorming process indeed indicate a larger degree of satisfaction and enjoyment from the experience, compared to nominal groups.

2. *Group experience*: Working in a group causes members to feel that something new was created in the course of the brainstorming session. To justify the efforts invested in the group process, individuals tend to believe that the group produced something in which all members had a part. In their evaluation of group outcomes, participating individuals do not distinguish between their ideas and the ideas of others. They are usually under the impression that they were the source of more ideas than was actually the situation. This misperception leads to an evaluation of an effective and satisfying process. In addition, owing to the high attentiveness, group members internalize the ideas of their fellow members, creating a false belief that more ideas were generated than would have been generated if each worked alone.

Other factors contributing to the illusion of the effectiveness of the brainstorming process cannot be identified and manipulated under laboratory conditions. A major reason for the method's popularity is found in the organizational functions served by the process, and the organizational benefits that are its result (Sutton and Hargadon, 1996).

- *Support of common organizational memory*: Brainstorming sessions help organization members acquire, store, retrieve, modify, and combine knowledge of various solutions to the problems they face. The sessions create opportunities to add new knowledge and solutions to the organizational memory. Furthermore, such sessions serve as an efficient means of distributing knowledge among organization members, reinforcing the knowledge of older members, and imparting organizational knowledge to new members, including solutions generated previously.
- *Diversification of ability*: Participants in brainstorming frequently define it as a pleasant and fun experience. Part of the enjoyment is related to the possibility of working with others in an unrestricted manner. For most participants, the session is a social encounter, an opportunity to share experiences and discharge everyday stress. Another benefit is the opportunity to use and learn new technologies and knowledge. In other words, brainstorming affords participants the opportunity to experience diversity and interest that is not always present in their everyday functions.
- *Competition over status*: Brainstorming is an important organizational arena in which competition over status takes place, based on the competencies of the group members, who meet for a predetermined session or sessions to concentrate on a specific problem. The rules of the brainstorming session are known to all: although bad ideas are not criticized, good ideas are praised, creating an opportunity to receive less threatening feedback from organization members.
- *Impressing clients*: Brainstorming is an opportunity for an organization to convey the competencies of its members. Quite often, clients are impressed by creativity expressed in meetings and they love the positive atmosphere. Organizations use this process to show their client that they understand his or her problem, and that they rely on a wealth of experience to arrive at the best solution. Brainstorming is also an efficient forum for the client to present his or her problems and gather ideas from a number of participants simultaneously.

Moreover, the statement "everybody uses brainstorming" is not necessarily entirely correct. First of all, a great deal of what is called *brainstorming* is no more than efficient teamwork. Let us recall an example we are all familiar with: students studying together for an exam. Not only are they not "brainstorming"

or raising far-fetched ideas, but what they are doing is actually the opposite. They solve the problems by focusing on each problem at hand. Imagine a Harvard Business School case analysis or solving a complex mathematical function—do students raise random ideas? Do they raise ideas that are detached from each other? Do they defer judgment? Of course, they do not. Even engineers who meet to solve problems do not conduct brainstorming sessions, though they may call it that. In practice, in most of their meetings, they conduct efficient discussions to examine various alternatives, and assist each other in solving the problem. In such a process, the group effect stems from focusing rather than from brainstorming.

The physicist Tom Hirschfeld once said "The second attack on the same problem should be from a completely different angle." Apparently, the group encounter does have a value that is not necessarily relevant to the mechanism of brainstorming. This can be illustrated by the story of the company that decided to eliminate its coffee corner. Exhibiting great efficiency, the company appointed an individual to bring coffee and cookies to anyone who sends a request through the company's Intranet. The company executive thought that by eliminating the need for the company's engineers to leave their desks, their efficiency would be improved, and he or she could also convey the consideration that the company shows its employees by affording them a higher degree of convenience. After a number of months, the executive found that eliminating the random meetings at the coffee machine prevented an exchange of knowledge and opinions, had a detrimental effect on the engineers, and their productivity fell.

This story indicates a possible means of reconciling our conflicting data. Perhaps the brainstorming method of discussion is not particularly effective for the generation of new ideas, but apparently it, or the encounter it creates, exposes previously untapped ideas of the organization's members. Imagine that an organizational member is facing a problem similar to the one previously solved by another member on the floor below. In an encounter such as brainstorming, he or she would be able to share his or her accumulated knowledge, the successful and failed tests that he or she conducted, and perhaps the directions that he or she examined and found potentially beneficial.

If so, how can we reap the benefits of brainstorming while avoiding its relative shortcomings, as found in laboratory experiments? In this context, we would like to present some new insights relating to brainstorming and suggest how a more advanced version of this approach may be used to the benefit of organizations. In electronic brainstorming (EBS), a recently introduced electronic version of brainstorming (Gallupe *et al.*, 1992), brainstorming sessions—rather than taking place aloud and in a single location—are conducted by virtually merging nominal groups and opening a new channel of idea-sharing and knowledge transference. Each organization member sits at his or her own desk, electronically connected to the others (sometimes from other plants or organizations). Participants generate ideas on their own and send them to the general pool, while continuing to develop their train of thought and generate more ideas on the theme. When they feel that they are ready to investigate the ideas of others, they download them from the pool. These inputs lead to new responses on their part or to further development or generation of ideas. This virtual neuron storm has a number of advantages: individual members can control their retrieval of the ideas of other participants and choose the timing for reflection on these inputs, thus avoiding the distraction effect. Outside observers can add new instructions or guidelines. In a global brainstorming process that was conducted through the Internet at the initiation of a well-known food company, corporate executives observed the exchange of ideas in real time, on a giant screen. They concurrently conducted a discussion to assess and filter the various ideas and even made real-time decisions.

Empirical findings reflect the success of EBS, both absolutely and relative to "regular" brainstorming. Large groups of participants generated more ideas than individuals working alone (in nominal groups). In addition, the quality of the EBS-generated ideas was rated as higher. The size of the EBS group enables a quick and sharp compilation of the conceptual capital in the organization. Apparently, the absence of inhibiting social phenomena that characterize

group processes during brainstorming and the ability to concentrate and develop a train of thought without interruption, avoid the pitfalls of brainstorming while enhancing its benefits. Although it is too early to claim that advanced EBS has made significant inroads into managerial practice at the cost of traditional brainstorming, its superiority has been sufficiently demonstrated to allow us here to seriously recommend its integration in the problem-solving modes of organizations. Nevertheless, we should remember that despite EBS' superiority in idea generation, it lacks the social encounter that has been found to be an important side-benefit of conventional brainstorming encounters.

Another interesting finding in regard to recent creativity studies (not conducted in the context of group research) relates to the perception that the constraints of a problem encourage the production of more creative ideas (Smith, Ward, and Finke, 1992). Constraints provide a focus for the brainstorming session and transform it into "focused storming." Brainstorming synergies are not a result of distractions. They are based on a number of minds working in a well-defined direction. In this context, we can mention that studies of group decision-making process indicate that when a discussion is well managed according to a decision-making model, the group has a higher value than the sum of its members. Findings such as these lead us to hypothesize that when the problem is constrained and well defined, even subject to complete freedom of expression, anarchy is not a possible outcome. In our opinion, a brainstorming session should be conducted to generate solutions to well-defined problems, with a clear set of criteria for success. In this case, information exchange, conceptual capital, and directing the storm of ideas to a target-focused channel contribute to organizational performance. Conversely, when problems are ambiguous and ill defined, such as problems relating to new-product innovation, focused storming is preferred.

To make the most of brainstorming, it is imperative to define the problem at hand and the goals of the encounter beforehand and in a precise manner. Brainstorming is one of the easiest methods to implement, and although many different ways of doing this are available, these should be consistent with the context and should fit the situation. It is important to remember that brainstorming is not always the only or the preferred option.

Bibliography

Diehl, M. and Stroebe, W. (1987) Productivity loss in brainstorming groups: toward the solution of a riddle. *Journal of Personality and Social Psychology*, **53** (3), 497–509.

Diehl, M. and Stroebe, W. (1991) Productivity loss in idea-generating groups: tracking down the blocking effect. *Journal of Personality and Social Psychology*, **61** (3), 392–403.

Gallupe, R.B., Dennis, A.R., Cooper, W.H. *et al.* (1992) Electronic brainstorming and group size. *The Academy of Management Journal*, **35** (2), 350–369.

Osborn, A.F. (1957) *Applied Imagination; Principles and Procedures of Creative Thinking* (Rev. ed.), Scribner, New York.

Paulus, P.B., Dzindolet, M.T., Poletes, G., and Camacho, L.M. (1993) Perception of performance in group brainstorming: the illusion of group productivity. *Personality and Social Psychology Bulletin*, **19** (1), 78–89.

Smith, S.M., Ward, T.B., and Finke, R.A. (1992) *Creative Cognition: Theory, Research, and Applications*, MIT Press.

Sutton, R.I. and Hargadon, A. (1996) Brainstorming groups in context: effectiveness in a product design firm. *Administrative Science Quarterly*, **41** (4), 685–718.

concept testing

Kenneth B. Kahn

Concept testing represents those activities undertaken during the concept evaluation phase of the new-product development process for the purpose of assessing and incorporating customer opinion before committing substantial funds (*see also* PRODUCT TESTING) (Moore, 1982). During concept testing, general consumers, current customers, and/or potential customers are asked to evaluate a new product concept by giving their opinion on whether the concept is something that they may have interest in and would likely buy. Other questions to be asked during a concept test can include the following: Why is the concept (or not) appealing? What are the main strengths of the product concept? What are the key weaknesses of the product concept? What are the suggestions for improving the proposed product concept (Cooper, 1993)? Focus groups, in-person interviews, telephone interviews, mail surveys, and online surveys are common data collection methods in the course of concept testing (Peng and Finn, 2008).

Types of Concept Testing

There are four general types or approaches for conducting a concept test: the narrative concept test, the pictorial concept test, the prototype concept test, and the virtual concept test. The narrative concept test involves a text description of the concept being presented to consumers for their opinions. There are two styles in which the text description or concept statement can be written: noncommercialized and commercialized. A noncommercialized concept statement, also called a *stripped description*, is a simple factual statement about a proposed product concept that contains a minimum of attributes; a commercialized concept statement, also called an *embellished description*, is a persuasive statement about a product concept written in such a way as to promote the product and its attributes similar to how a product would be promoted at launch. Lees and Wright (2004) find marginal differences when comparing the attitude and purchase intention responses to the two concept statement styles.

Instead of text, the pictorial concept test involves a black and white or color drawing being presented to consumers for their opinions. The prototype concept test involves the customer being shown and possibly being asked to handle a facsimile of the product and/or an actual working version of the product (*see* PROTOTYPE). The benefit of a prototype is that the consumer can witness and experience the product firsthand. The virtual reality concept test is when the consumer views a computer image of the product and possibly interacts with a virtual prototype, thereby virtually witnessing and experiencing the product.

While there are four types of concept tests, product concepts are typically evaluated through a combination of these concept tests. For example, a concept test can be staged to observe consumers' evolving opinions as a result of being first exposed to a narrative text statement (narrative concept test), then to a picture (pictorial concept test), and finally to a working prototype (prototype concept test). Each subsequent test provides additional information that the consumer can process and respond to. In choosing between a narrative concept test and a pictorial concept test, one should consider whether the research objective is to show concept test participants how the product looks so that they can gauge an opinion, or whether it is better to have concept test participants attempt to visualize what the product would/should look like. The latter may be useful in designing an idealized version of the product according to consumer opinion.

Bibliography

Cooper, R.G. (1993) *Winning at New Products: Accelerating the Process from Idea to Launch*, 2nd edn, Addison-Wesley, Reading.

Crawford, C.M. (1997) *New Products Management*, 2nd edn Irwin, Boston.

Dolan, R.J. (1993) Concept testing, *Managing the New Product Development Process*, Addison-Wesley Publishing Company, Reading, pp. 83–93.

Lees, G. and Wright, M. (2004) The effect of concept formulation on concept test scores. *Journal of Product Innovation Management*, **21** (6), 389–400.

Moore, W.L. (1982) Concept testing. *Journal of Business Research*, **10** (3), 279–294.

Peng, L. and Finn, A. (2008) Concept testing: the state of contemporary practice. *Marketing Intelligence and Planning*, **26** (6), 649–674.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

product testing

C. Anthony Di Benedetto

When a product is in the technical development stage, the product team must ensure that it is on track to meet objectives. In earlier stages, before a physical prototype is available, the team would likely have conducted concept tests with potential customers (*see* CONCEPT TESTING). Now, however, at least an early prototype is available for customer trial, and valuable information may be obtained from customers as to their initial reaction to the physical prototype, their level of satisfaction with the prototype, and also their likes and dislikes. This information can be used to further refine the prototype as well as to assist in making other marketing decisions, such as brand name, package design, and price levels. We begin with a brief discussion of prototypes and then move to the objectives and methods of product testing.

Prototype Development[1]

There are different levels of prototypes, each used at different stages in product testing. The term *focused prototype* refers to a nonfunctioning or limited-functionality prototype, which product developers use to assess early customer reactions. After further development, a "comprehensive prototype," or a basically complete version, is prepared and can also be taken to market to make final adjustments to the product and supporting elements such as price or brand name. A cell-phone manufacturer may make several nonfunctioning phones out of wood or foam, and see whether customers like the size, how it feels in the hand, how big the screen is, how close together the buttons are, and can make appropriate changes. Many of these can be made at very low cost, until the firm has zeroed in on customer likes while eliminating the dislikes. In fact, the more radical the new product is, the more the manufacturer may want to use a series of focused prototypes to better understand elusive customer requirements. This process is sometimes called *probe and learn*. For example, Iomega reportedly went rapidly through about 50 nonfunctional prototypes of their new high-storage disk drive, which eventually became known as the *Zip Drive*, before settling on one that customers liked (Lynn and Reilly, 2002). Similarly, Charger Electric Bicycles built a succession of nonfunctioning bicycles out of wood and plastic to assess preferences for appearance, performance, and weight. They eventually also built a crude functional prototype (i.e., not beautiful, but it did work well enough to get customer trial reactions) and continued to make improvements to it. Eventually, all the technical and cosmetic trade-offs were made and the bicycle was commercialized (Mascitelli, 2000).

Of course, prototypes can also be developed using computer-aided design and manufacturing (CAD/CAM) techniques (*see* PRODUCT DESIGN). Using a technique sometimes called *rapid prototyping*, a three-dimensional computer model is converted into a hard plastic prototype, which can be evaluated by customers. If customers like it, the prototype can become the "master" from which the molds for the manufacturing process can be built (Van Dierdonck, 1990). Car companies routinely design cars using CAD/CAM and can execute thousands of virtual "crash tests," thus reducing the need for expensive physical prototype manufacture (Thomke, 1998; Thomke and Fujimoto, 2000). Nevertheless, some car companies such as BMW will complement CAD/CAM car design with traditional, painstaking clay models, to get the minute design features just right (Bangle, 2001). Fixing these problems at this early stage minimizes the need for expensive retooling later in the process.

Product-Use Testing

Now that the product prototype has been built, we can proceed to product-use testing. This step is sometimes called *beta testing*, a term that originated in the computer software industry (alpha testing means testing with internal employees, beta testing with outsiders). Customer needs can be very complex, and it may be necessary to conduct multiple product-use tests to come up with a successful version of the product, especially in the case of a radically new product. General Electric, for example, tested four different versions of its computed tomography (CT) scanner – each offering

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

different functionality to the physician – before hitting upon the full-body scanner that became the marketplace success (Lynn *et al.*, 1998). Careful product-use testing will reveal problems that may be easy to rectify but might otherwise not be identified until it is too late. According to the web site www.baddesigns.com, Polaroid once launched a children's camera that could be opened (thus ruining the film inside) if the child simply pressed a button (Crawford and Di Benedetto, 2008, p. 352). Somewhat unexpectedly, the problem was that the camera was too easy to open! Better product-use testing might have found and corrected this problem.

Knowledge gained from product-use testing[2]. There are four categories of knowledge that can be obtained from the product-use test: preuse sense reactions, early use experiences, major benefit results, and diagnostic information.

As the name implies, preuse sense reactions are customer sensations of the product even before they have a chance to evaluate it. A carmaker will realize that the clean, inviting appearance of the showroom, and the helpfulness and expertise of the sales staff, will make a good first impression and help in creating a positive customer sensation. Alternatively, a product may make a promise that customers will perceive as "too good to be true." A new headache remedy that is both very gentle and very effective might meet customer skepticism, despite the fact that it really works. The product-use test is an opportunity to assess the believability of the claim, and to take corrective action. For example, at the time of the iPhone launch if the average person would find it hard to believe that it really is a music player, photo album, and internet device built into a phone, then perhaps demonstration promotional campaigns are required in order to show the product in use.

Early use experiences are also assessed at this stage. The first few features that the user will be able to judge will be, for example, whether the product is easy to use, whether there are any apparent problems while using it, whether the product seems to do what it is supposed to do, and any other early impression.

The heart of the product-use test is the assessment of major benefits (*see* VALUE PROPOSITION). In a traditional beta test of, say, a new piece of banking software, the developer will get the product adopted at a given number of customer sites, and the product is put to use. Users may be given checklists to assess specific benefits that should be provided by the new software (such as whether it is performs as advertised, whether it is "robust" to mistaken entries, whether it provides cost benefits, etc.). Notice that the goal of the product-use test is not necessarily to test specifically if the product meets customer needs, but rather to make sure it provides the benefits it is supposed to.

While traditional product-use testing involves a small number of customer sites, this is changing, notably in computer software. Most of us have downloaded free beta versions of new software, with the usual disclaimers. When a software developer releases beta versions for download, there may be thousands of customers who download and try the software. In a classic case, both Netscape and Microsoft were developing new Internet software in 1996. During this period, both firms were releasing limited-functionality alpha versions on company web sites to employees, followed by multiple beta versions to customers. Needless to say, Microsoft employees were downloading Netscape beta versions and vice versa, so as to keep track of competitive developments. During this period, Microsoft went from having virtually no presence in this market to having a fully developed early commercial release of the Internet Explorer (Iansiti and MacCormack, 1997).

Finally, the manufacturer also seeks to obtain diagnostic information from the product-use test. The users at the customer site may suggest performance improvements, or verify whether the product's delivered benefits match the intended positioning of the product. Alternatively, warning signs may be identified: customers may report that the new product is hard to understand or they found the presentation of results difficult to interpret.

It should also be noted that some firms do not go through all of these steps in such a linear fashion. Some managers feel that they get sufficient marketplace information with a minimum of testing. As an example, during the period when it managed the Snapple line of drinks, Triarc managers dropped all product-use testing, relying on their own instincts for new

flavors and trying these themselves – in short, they conducted small-scale alpha testing only, followed immediately by a launch. The management, no doubt, felt that the downside risks of launching a weak product were not too steep, and indeed had success with this method; the popular Snapple element brands (Rain, Sun, and Fire) were launched during this time (Deighton, 2002).

Product-use testing considerations. The first rule for effective product-use testing is that the product being tested should be as similar as possible to the product that will eventually be manufactured and marketed. Manufacturers have been known to product-use test a pilot version of the product that is far higher in quality than the mass-produced product ultimately launched. Related to this, if there is any possibility of customer misuse of the product, the product should be tested by real consumers in real situations, rather than under exacting laboratory conditions that would never exist in reality. A new instant coffee that tastes perfect when made with distilled water and spooned out with chemical measuring instruments might be terrible when carelessly prepared (for more similar misuses of product-use testing, see Lavidge, 1984).

There are other guidelines as well. The product developer must carefully specify the performance requirements to be used by the customer site. For example, a term like *user-friendly* can mean different things to different people, so unless it is clearly defined, it is open to multiple interpretations, rendering the results meaningless. In addition, timing is also a factor. Running the product-use test too late in the development process guarantees that time to market will be adversely affected, and delays will occur, if the developer tries to fix problems uncovered during the test. However, if the test is conducted too early, there may be too many bugs still in the product and some bad word of mouth publicity may take place. Finally, if the product-use test results are poor, they should not be ignored. All results from the product-use test are valuable and should be evaluated and used to improve the product, regardless of how negative they are (Stoy, 1996).

Endnotes

[1] Parts of this section are derived from Ulrich and Eppinger (2000), Chapter 12.

[2] Parts of this section are derived from Crawford and Di Benedetto (2008), Chapter 15.

Bibliography

Bangle, C. (2001) The ultimate creativity machine: how BMW turns art into profit. *Harvard Business Review*, **79** (1), 47–55.

Crawford, M. and Di Benedetto, A. (2008) *New Products Management*, 9th edn, McGraw-Hill/Irwin, Burr Ridge.

Deighton, J. (2002) How Snapple got its juice back. *Harvard Business Review*, **80** (1), 47–53.

Iansiti, M. and MacCormack, A. (1997) Developing products on Internet time. *Harvard Business Review*, **75** (1), 108–117.

Lavidge, R.G. (1984) Nine tested ways to mislead product planners. *Journal of Product Innovation Management*, **1** (2), 101–105.

Lynn, G.S., Mazzuca, M., and Morone, J.G. (1998) Learning is the critical success factor in developing truly new products. *Research-Technology Management*, **41** (3) 45–51.

Lynn, G.S. and Reilly, R.R. (2002) *Blockbusters: The Five Keys to Developing Great New Products*, Harper Collins, New York.

Mascitelli, R. (2000) From experience: harnessing tacit knowledge to achieve breakthrough innovation. *Journal of Product Innovation Management*, **17** (3), 187–188.

Stoy, R. (1996) Assembled product development, in *The PDMA Handbook of New Product Development* (eds M.D. Rosenau, A. Griffin, G. Castellion, and N. Anscheutz), John Wiley & Sons, Inc., New York, pp. 271–286.

Thomke, S. (1998) Simulation, learning, and R&D performance: evidence from automotive development. *Research Policy*, **27**, 55–74.

Thomke, S. and Fujimoto, T. (2000) The effect of 'front-loading' problem-solving on product development performance. *Journal of Product Innovation Management*, **17** (2), 110–127.

Ulrich, K. and Eppinger, S.D. (2000) *Product Design and Development*, 2nd edn, Irwin/McGraw-Hill, Burr Ridge.

Van Dierdonck, R. (1990) The manufacturing-design interface. *R&D Management*, **3**, 203–209.

core competencies

Shikhar Sarin

Introduction

In their highly influential book "Competing for the Future," Hamel and Prahalad (1994) suggest that in the short run, large diversified firms derive their competitiveness from the price and performance attributes of their current products. However, over time, competition forces most firms to converge on similar standards of cost and quality, eroding any differential competitive advantage. In the long run, the real source of competitive advantage lies in the ability of the senior managers to consolidate organization-wide technologies and skills into competencies that create unanticipated products; competencies that allow firms to build products faster and at a lower cost than competitors, and to adapt nimbly to changing environments (Hamel and Prahalad, 1994; Prahalad and Hamel, 1990).

Definition

Core competencies are collective learnings in an organization that result from a unique coordination of diverse production skills and/or the integration of multiple streams of technologies (Prahalad and Hamel, 1990). Unlike physical assets, core competencies do not diminish with use, nor do they deteriorate over time; instead they are enhanced as they are applied and used (Hamel and Prahalad, 1994).

Competing for the Future

Identifying core competencies. The notion of core competencies has been criticized for being too abstract, ill defined, and difficult to implement. Managers often have a hard time identifying core competencies of their firms. However, Prahalad and Hamel (1990) offer three tests that can be applied for identifying core competencies in an organization:

- Core competencies provide potential access to a wide variety of markets.
- Core competencies make a significant contribution to perceived customer benefits of the end product.
- Core competencies are hard for competitors to imitate.

Hamel and Prahalad (1994) suggest that few firms are likely to be world leaders in more than five or six fundamental competencies. Those that name a laundry list of competencies are really not identifying their *core competencies* as subject to the three tests listed above.

Roots of competitive advantage. Battles for global leadership are waged on three different planes: core competence, core product, and end products (Hamel and Prahalad, 1994). It is critical for managers to understand and appreciate the distinction between these levels. Core products are physical manifestations of one or more core competencies. They are the tangible link between core competencies and end products. Core products are components or subassemblies that contribute to the value of multiple end products. For example, Honda's core competence in design and development manifests itself in the form of the global leadership in engines and power trains. These engines and power trains then lead to a proliferation of end products for Honda in markets as diverse as cars, motor cycles, generators, lawn mowers, compressors, and so on (*see also* PRODUCT PLATFORMS).

Prahalad and Hamel (1990) suggest that one way to distinguish between a core competence, core product, and end product is to visualize a diversified firm as a tree. The trunk and major limbs of the tree represent the core products; the smaller branches are the strategic business units (SBUs) (*see* ORGANIZING FOR INNOVATION); the leaves, flowers, and fruits represent the end products. The root system of this tree, which provides its underlying strength, nourishment, stability, and sustenance, is the core competence. While competitive advantage in end products is relatively short-lived, nurturing of core competencies and core products can result in long-term and sustainable competitive advantage.

Rethinking the corporation for the future. Most large diversified firms are organized as portfolios

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

of relatively independent and autonomous SBUs. Owing to the primacy of the SBU structure, resources, technologies, talent, and knowledge become trapped in isolated pockets within large organizations, with few opportunities for cross-fertilization. Moreover, in large diversified organizations, no entity is responsible for developing a comprehensive strategy that transcends traditional SBU boundaries. Because of imprisoned resources, innovation becomes bounded and is limited to marginal line and geographic extensions. This results in a chronic underinvestment in, and underdevelopment of, core competencies and core products (Prahalad and Hamel, 1990).

Managers of diversified organizations spend a disproportionate amount of their time and energy in competing over end products, while overlooking the far more critical levels of competition in core products and core competencies (Hamel and Prahalad, 1994). To claim global leadership, executives of diversified firms will need to break free from the "tyranny of the SBUs," and start regarding the organizations as portfolios of core competencies and core products, rather than as a collection of discrete businesses.

Bibliography

Hamel, G. and Prahalad, C.K. (1994) *Competing for the Future: Breakthrough Strategies for Seizing Control of Your Industry and Creating Markets of Tomorrow*, Harvard Business School Press, Boston.

Prahalad, C.K. and Hamel, G. (1990) The core competence of the corporation. *Harvard Business Review*, **68** (3), 79–91.

competitive advantage

Shikhar Sarin

Introduction

Two key issues guide the choice of competitive strategy for a firm: attractiveness of the industries for long-term profitability, and the determinants of relative competitive position within an industry (Porter, 1985). The ensuing discussion focuses on the latter.

Definition

Competitive advantage is a firm's ability to grow in spite of competition. Best (2004, p. 387) defines *competitive advantage* as the relative advantage one business has over another that is sustainable and translates into a benefit that is important to target customers. There are two basic types of competitive advantages: cost leadership and differentiation (Porter, 1985). Cost leadership stems from the ability to offer lower prices than competitors for equivalent benefits; whereas, differentiation stems from the firm's ability to provide unique benefits that more than offset the premium price charged by the firm. Attaining competitive advantage requires a firm to make choices regarding what kind of advantage it seeks, and the scope within which the firm seeks to achieve it. Porter (1985) contends that combined with the choice of competitive scope (or the range of a firm's activities), these two types of competitive advantages can lead firms to three different kinds of generic business strategies for achieving above-average profitability.

Generic Business Strategies

Cost leadership. In this type of strategy, a firm sets out to become *the* low-cost producer in its industry by selling a standard, no-frills product (Porter, 1985). The firm operates with a broad scope and serves many industry segments. The cost advantage of the firm can stem from many different sources such as economies of scale, control of proprietary technology, preferential access to resources, and so on. However, a cost leader must achieve parity or proximity in the basis of differentiation relative to its competition. In other words, the cost leader should either offer an identical product offering to competitors, at a lower price, or a different combination of product attributes that is equally preferred by buyers (*see also* VALUE PROPOSITION). A successful implementation of this strategy requires that the firm be *the only* cost leader in an industry. If there is more than one firm competing on this strategy, it can lead to disastrous consequences for industry structure and overall profitability.

Differentiation. Under this strategy, the firm selects one or two dimensions that the buyers value and uniquely positions itself to meet those needs – while charging a premium price for such uniqueness (*see* PRODUCT POSITIONING). The bases for differentiation can arise from several sources, such as the product itself, the method of delivery/distribution, and the marketing approach. A differentiator must achieve cost parity or proximity with its competitors by reducing cost in areas that do not affect differentiation; otherwise its price premium will be neutralized by inferior cost structures (Porter, 1985). Porter suggests that since differentiation is based on a firm distinguishing itself on attributes that are different from its competitors, there can be more than one successful differentiator in an industry.

Focus. This strategy chooses a narrow competitive scope within an industry (i.e., a niche). The focuser selects a segment or a group of segments in the industry, and seeks to achieve competitive advantage in the target segment(s). In other words, the focuser optimizes its strategy on a narrow scope of activities, even though it may not enjoy a competitive advantage overall. This strategy rests on differences between the target segment(s) and other segments in the industry. The target segment must have unusual needs or preferences that are poorly served by broadly targeted competitors. The focuser can exploit these differences and achieve competitive advantage by dedicating itself to serving these segments exclusively. Porter (1985) suggests that there are two variants to this strategy: differentiation focus and cost focus. In differentiation focus, the firm seeks differentiation in its target segment(s) by exploiting the special needs of

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

buyers in the target segment(s). Cost focus, on the other hand, exploits the differences in cost behavior of the target segment(s). Porter suggests that there is room for several sustainable focus strategies in an industry, provided that focusers choose different target segments.

Sustainability of Competitive Advantage

Sustainability of these generic strategies demands that a firm's competitive advantage resist erosion by competitive action or industry evolution. Porter (1985) suggests that cost leadership is not sustained if competitors imitate, technology changes, or bases for cost leadership erode. Cost leadership is also hard to sustain if the cost leader loses proximity in differentiation, or cost focusers achieve even lower costs in segments. Similarly, differentiation is not sustainable if competitors imitate, or bases for differentiation erode. Differentiation strategy is at risk if the differentiator loses cost proximity, or differentiation focusers achieve even greater differentiation in segments. In addition to dangers of imitation, focus strategy is also at risk if the target segment(s) become structurally unattractive or the demand disappears. Broadly targeted competitors can overwhelm the focusers if the target-segment differences from other segments narrow down, or the advantages of a broad line increase. Focusers can also be at risk if new focusers further subsegment the industry.

Ghemawat (1986) notes that competitive advantage can be sustained by not ignoring any contestable advantage that may crop up. He argues that sustainability is greatest when it is based on several kinds of advantages rather than just one. To create a sustainable competitive advantage, one must be able to preempt competitors, especially in industries that offer potent first-mover advantage (Ghemawat, 1986; *see also* FIRST-MOVER ADVANTAGE). In addition, competitive advantage in one industry can be enhanced by interrelationships with business units in other industries (Porter, 1985; *see also* CORE COMPETENCIES).

Bibliography

Best, R.J. (2004) *Market-Based Management: Strategies for Growing Customer Value and Profitability*, 3rd edn, Prentice Hall, Upper Saddle River.

Ghemawat, P. (1986) Sustainable advantage. *Harvard Business Review*, **64** (5), 53–58.

Porter, M.E. (1985) *Competitive Advantage: Creating and Sustaining Superior Performance*, The Free Press, New York.

locale of innovation

K. Sivakumar

Introduction and Definitions

For the purpose of this discussion, "local" means involving only one country. "Transnational" means involving more than one country and, typically, several countries. In that sense, "transnational" is used synonymously with "global," "multinational," "international," "cross-country," "multicountry," and similar terms. "Innovation" can be conceptualized as (i) the process of developing new products (Crawford and Di Benedetto, 2006; Belliveau, Griffin, and Somermeyer, 2002) or (ii) the new product outcome itself (*see* INNOVATION TYPOLOGIES). "Product" is used as a generic term to include goods, services, ideas, and processes, including both radical and incremental products (*see* GLOBAL PRODUCT DEVELOPMENT). Combining the two concepts, the phrase "local innovation" can mean that (i) the development of the innovation is locally done or (ii) the innovation caters to the local markets; similarly, the phrase "transnational innovation" means that (i) the development of the innovation is transnational or (ii) the innovation can cater to transnational consumption. The following typology helps us understand the entire gamut of meanings of these concepts.

		Locus of development	
		Local	*Transnational*
Locus of consumption	Local		
	Transnational		

Scope and Meaning of Local and Transnational Innovation

Although most innovations are associated with business entities, it must be realized that innovation can be a social phenomenon and can be developed exclusively by companies, partially by companies in association with universities and/or communities, or entirely by universities and/or communities.

With respect to innovations in organizations and the various natures and extents of globalization (e.g., from a small manufacturer supplying parts to a US company to truly multinational companies such as IBM with operations and suppliers in several countries and diversified companies with different kinds of businesses scattered all over the world), the phrases "local innovation" and "transnational innovation" must be understood in their proper context. For example, a local innovation process involving several divisions in a large company like General Electric may involve more complexities than a transnational innovation process of new software development between two different companies in two different countries or two teams of the same company located in two different countries (*see* CROSS-FUNCTIONAL TEAM; VIRTUAL TEAMS). If a local innovation is an incremental innovation, the original innovation, to which minor modifications are made, could have been a transnational innovation or a local innovation; a local innovation might also simply be a product that is modified locally to cater to local needs (Bartlett, Doz, and Hedlund, 1990; Bartlett and Ghoshal, 1998).

Similarly, when discussing the process of transnational development of innovation (Eppinger and Chitkara, 2006; Nakata and Sivakumar, 1996), the original innovation (e.g., a patent for the discovery of a molecule) could have been developed locally in one country, but subsequent innovation (e.g., drug development using the patent to develop medicines to cure illnesses) may be transnational. Another example is a multinational company that uses all its resources in multiple countries to develop a water-purification system (i.e., a transnational innovation process) to cater to the specific water composition in an African village (to be consumed as a local innovation).

When we move away from innovations that companies spearhead into the realm of societal changes (e.g., modification of a teaching method to address the lack of technology in a developing country's village), a transnational innovation is modified to become a local innovation. Sometimes local innovation also expands to include regions. For example, the concept of microcredit can be considered a local innovation developed in a single developing country but can become

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

a transnational innovation if several countries in a given region and beyond adopt it. Similarly, systems to organize rural farmers into powerful cooperatives can be local when based on what is done locally in a particular country (e.g., India) but also can become transitional (e.g., Pakistan and Sri Lanka).

Another aspect to consider is the production of the innovation – the innovation can be locally or transnationally developed for local or transnational consumption and can be produced in each local market or in transnational facilities. In this case, the two-by-two matrix developed earlier could be converted into a three-dimensional matrix to include development, production, and consumption. Interestingly, the production process may involve local modifications to the innovation (i.e., local innovation), thereby further developing incremental innovations.

Conclusion

As the movement of people, products, and ideas across countries is increasingly becoming a norm for business and nonbusiness organizations, in particular, and society at large, in general, innovation activities and the outcomes of such activities are becoming increasingly transnational (Cooper, 2009). While purely local innovations will become more of a rarity, transnational innovations will involve different levels of involvement by people and organizations across different countries and/or firms.

Bibliography

Bartlett, C.A., Doz, Y., and Hedlund, G. (1990) *Managing the Global Firm*, Routledge, London.

Bartlett, C.A. and Ghoshal, S. (eds) (1998) *Managing Across Borders: The Transnational Solution*, 2nd edn, Harvard Business School Press, Boston.

Belliveau, P., Griffin, A., and Somermeyer, S. (eds) (2002) *The PDMA Handbook for New Product Development*, John Wiley and Sons, Inc., New York.

Catwell, J. (ed.) (1994) *Transnational Corporations and Innovatory Activities*, Routledge, London.

Cooper, R.G. (2009) The state of product development. *Research-Technology Management*, **52** (1), 6–7.

Crawford, M. and Di Benedetto, A. (2006) *New Products Management*, 8th edn, McGraw-Hill Irwin, New York.

Eppinger, S.D. and Chitkara, A.R. (2006) The new practice of global product development. *MIT Sloan Management Review*, **47** (4), 22.

Friedman, T.L. (2005) *The World is Flat*, Farrar, Straus and Giroux, New York.

Golder, P.N. (2000) Insights from senior executives about innovation in international markets. *Journal of Product Innovation Management*, **17**, 326–340.

Nakata, C. and Sivakumar, K. (1996) National culture and new product development: an integrative review. *Journal of Marketing*, **60**, 61–72.

Söderquist, K.E. (2006) Organising knowledge management and dissemination in new product development: lessons from 12 global corporations. *Long Range Planning*, **39** (5), 497.

integrated product development

Fred Langerak

Integrated new-product development is a contemporary managerial approach for improving new-product development performance through the overlap (partial or complete parallel execution) of and the interaction (exchange of information) between activities in the new-product development process (Naveh, 2005). Since overlap and interaction increase the need to coordinate activities, integrated product development requires the systematic employment of teams consisting of members from all relevant disciplines to integrate and concurrently carry out the critical activities to efficiently develop a new product that satisfies customers' needs.

Teaming involves the establishment of cross-functional teams for the purpose of better integrating and coordinating activities between different organizational units, as well as coordinating the organization's interaction with other firms (*see* CROSS-FUNCTIONAL TEAM). The emphasis is on the early and active involvement of all internal (e.g., design, manufacturing, marketing) and external (e.g., suppliers and customers) stakeholders to create ideas and transform them into designs (*see* PRODUCT DESIGN), prototypes (*see* PROTOTYPE), and plans for product launch (*see* LAUNCH STRATEGIES) (Koufteros, Vonderembse, and Doll, 2002). In this way, problems that might emerge in the later stages of the development process are taken into consideration in the earlier stages when they can still be dealt with at a lower cost and without causing serious delays.

Among the benefits of integrated product development reported in the academic literature is reduced development time and reduced new-product goal failure. In the business community, integrated product development is also held to promote better knowledge sharing between different functional groups, enhanced flexibility, cost savings, clear focus on risks, and improved new-product quality (Nellore and Balachandra, 2001). Frequently reported drawbacks of integrated new-product development are lower task specialization and decreased project innovativeness, especially if it is applied in the early stages of a development project.

Integrated product development studies fall into four categories of research (Gerwin and Barrowman, 2002). The organizational design approach investigates integrated product development in terms of structural adaptations in new-product strategy and project organization and execution to environmental and technological changes. In the information-processing approach, integrated product development is viewed as a problem-solving process that allows an organization to generate, disseminate, and use information on market opportunities and technological possibilities, and translate the resulting knowledge into a product-design solution that meets customer requirements. Integrated product development is also studied as the application of total quality management principles from manufacturing to new-product development, with integrated product development concepts such as virtual incorporation of customers and suppliers into cross-functional teams having their roots in this line of research. Finally, integrated product development is researched from the practices that firms use to achieve integrated product development, such as quality function deployment, rapid prototyping, digital product modeling, and joint design reviews (Boyle, Kumar, and Kumar, 2006).

Integrated product development is used in about 65% of all development projects, albeit in various manifestations such as design for manufacturing, concurrent engineering, early customer involvement, and time-based competition, and as such has become the standard for managing product development (Griffin, 1997). Many researchers suggest that the effectiveness of integrated product development practices depends (heavily) on organizational factors and project characteristics, such as the organization's innovative climate, senior management support, project organization, and leadership (*see* LEADERSHIP ROLES IN PRODUCT DEVELOPMENT), the complexity of the organization's development activities, and the project's newness. However, there exists no empirical consensus on the precise influence of these contingencies on the impact of integrated product development on new-product development performance.

Regardless of the contingencies involved, it is clear that while integrated product development

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

efforts should be initiated, planned, and led from the top down, they should be executed from the bottom up to foster ownership among cross-functional team members. Team members should place emphasis on understanding customers' needs and the processes required to design and deliver a product solution, aligning development projects with business strategy, facilitating early involvement and parallel design, designing products and manufacturing and support processes in parallel, and involving customers and suppliers early in the development process.

Acknowledgment

The author thanks Pinar Cankurtaran for her insightful and helpful comments.

Bibliography

Boyle, T., Kumar, V., and Kumar, U. (2006) Determinants of integrated product development diffusion. *R&D Management*, **36** (1), 37–54.

Gerwin, D. and Barrowman, N.J. (2002) An evaluation of research on integrated product management. *Management Science*, **48** (7), 938–953.

Griffin, A. (1997) The effect of project and process characteristics on product development time. *Journal of Marketing Research*, **34** (1), 24–35.

Koufteros, X.A., Vonderembse, M.A., and Doll, W.J. (2002) Integrated product development practices and competitive capabilities: the effects of uncertainty, equivocality, and platform strategy. *Journal of Operations Management*, **20** (4), 331–355.

Naveh, E. (2005) The effect of integrated product development on efficiency and innovation. *International Journal of Production Research*, **43** (13), 2789–2808.

Nellore, R. and Balachandra, R. (2001) Factors influencing success in integrated product development (IPD) projects. *IEEE Transactions on Engineering Management*, **48** (2), 164–174.

product-line strategies

Jeffrey B. Schmidt

Product-line Defined

A product line consists of multiple products (or services) that fulfill the same or similar needs and wants. Normally, they function in a similar manner, and they may use similar technologies, be sold to the same customers, and use the same distribution channels. Product lines can be described by their breadth and depth. Various branding issues are also related to product-line decisions.

Product-line Breadth and Depth

Decisions about the breadth and depth of product lines are important. The breadth of a product line refers to the number of different lines an organization offers. As Figure 1 shows, product-line breath may be described as either narrow or broad. Nike's product line is broad since it offers shoes, apparel, hockey equipment, golf equipment, and other products. Conversely, Palm is an organization that offers a narrow line of products consisting mainly of smart phones, electronic organizers, and various accessories for these devices.

More important are decisions about the depth of a product line, which refers to the number of items in that line. Product-line depth may be described as shallow or deep. As an example, Apple has a shallow line of iPhones. There are two colors available (i.e., black or white) and two different storage capacities (8 or 16 GB) for a total of four different iPhone models. With respect to cars, BMW has a deep line of products in its "3 series" automobiles. There are at least 17 different 3-series models.

Product line depth	Product line breadth: Narrow	Product line breadth: Broad
Shallow	Few product lines with a few items in each	Many product lines with a few items in each
Deep	Few product lines with many items in each	Many product lines with many items in each

Figure 1 Product-line breadth and depth.

Branding Issues (Family Brand, Company Brand)

In addition to the breadth and depth of product lines, the branding strategy should be carefully considered as it is important for product growth strategies and launch strategies (*see* GROWTH STRATEGIES; LAUNCH STRATEGIES, respectively). Rao, Agarwal, and Dahlhoff (2004) extended the branding taxonomy of Laforet and Saunder's (1994), and categorized branding strategies into three types – corporate branding, house of brands, and mixed branding, which are defined in the following sections.

Corporate branding. Here the corporate name is dominant in endorsing all or part of the firm's product and service brands. At a minimum, the corporate name is an element of the product brand names. This holds throughout all its subsidiaries and at all company levels. Examples of companies that employ this strategy are Hewlett-Packard, Motorola, and Federal Express.

Mixed branding. Organizations typically employ a set of house or family brands, such as subsidiary names, in their brand portfolio, in addition to using the corporate name for certain products. Brands with names other than the organization's name are typically strong and significant to the organization. Pepsi is an example of such a firm because apart from its flagship brand, it operates with Mountain Dew and Aquafina, and its subsidiaries Tropicana and Frito-Lay use individual brands at the product level (e.g., Doritos and Ruffles).

House of brands. Here the organization does not use its corporate name or the name of its subsidiaries. Instead, it uses individual brand names to market its products. Companies such as Unilever and P&G keep their corporate name in the background and use individual brands for their product lines.

Table 1 provides examples of firms following each of these branding strategies and their selected brands.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Examples of branding strategies.

Corporate Branding	*Mixed Branding*		*House of Brands*	
Company and Brand	*Company*	*Selected Brands*	*Company*	*Selected Brands*
Porsche	Gillette	Gillette, Oral-B, Duracell, Braun, Waterman	P&G	Pampers, Crest, Tide, Ariel, Bounty, Always, Febreze
AT&T	The Gap	The Gap, Banana Republic, Old Navy	Darden Restaurants	Red Lobster, Olive Garden, Bahama Breeze
Dell	3M	3M, Scotchguard, Thinsulate, Scotch	Bristol-Myers Squibb	Clairol, Aussie, Herbal Essences, Viactiv, Boost

Adapted from Rao, Vithala R., Manoj K. Agarwal and Denise Dahlhoff (2004), "How is Manifest Branding Strategy Related to the Intangible Value of a Corporation?," Journal of Marketing, 68 (October), 126–141.

Bibliography

Harrell, G.D. (2009) *Concepts of Marketing*, Chicago Education Press, Chicago.

Laforet, S. and Saunders, J. (1994) Managing brand portfolios: how the leaders do it. *Journal of Advertising Research*, **34**, 64–76.

Rao, V.R., Agarwal, M.K., and Dahlhoff, D. (2004) How is manifest branding strategy related to the intangible value of a corporation? *Journal of Marketing*, **68**, 126–141.

GE/McKinsey matrix

Ashish Sood

The GE/McKinsey matrix is a nine-cell portfolio matrix tool developed by McKinsey & Company for GE's large portfolio of strategic business units (SBUs) (see Table 1) (*see also* PORTFOLIO MANAGEMENT). The aim of the portfolio analysis is to develop a strategy to identify SBUs that should be prioritized in investment and growth strategies, and to review SBUs that should no longer be retained. The two factors used in the analysis are the attractiveness of the relevant industry and the SBU's competitive strength within that industry.

The vertical axis is the industry attractiveness, which is determined by a combination of factors such as market growth rate, market size, industry profitability, industry rivalry, and global opportunities (*see also* GROWTH STRATEGIES). The horizontal axis is the strength of the business unit, which is determined by a combination of factors such as market share, growth rate, brand equity, channel equity, and profitability (*see also* MANAGING MATURE PRODUCTS). Managers may also include a different set of factors including the relative importance of each factor to determine a quantitative measure of each axis.

The nine cells in the matrix can be grouped into three major segments. Cells 1–3 comprise the first segment, which includes strong business in an attractive industry. Cells 4–6 comprise the second segment, which includes SBUs with medium attractiveness either because the firm has strong SBUs in an unattractive market or weak SBUs in an attractive industry. Cells 7–9 comprise the third segment, which includes weak SBUs in unattractive markets. Managers can formulate appropriate strategies to strengthen their positions in the first segment, to move from the second segment to the first, and to divest SBUs in the third segment (*see also* PRODUCT DIVERSIFICATION).

The matrix may also be plotted graphically. The two axes are industry attractiveness and business unit attractiveness respectively. Each SBU is displayed as a circle. The size of the circles represents the market size. Market share is shown by using the circle as a pie chart and the direction and the movement of the SBUs in the future is shown by an arrow.

The GE/McKinsey matrix was designed as an extension of the BCG growth-share matrix (Henderson, 1970). The GE/McKinsey matrix has nine cells in contrast to the BCG Matrix that has only four and enables more detailed analyses. Some limitations of this matrix are that it is too complicated and too static, and it neglects the interactions with the environment and among SBUs.

Table 1 GE/McKinsey Matrix.

		Business Unit Attractiveness		
		High	*Medium*	*Low*
Industry Attractiveness	High	1	3	6
	Medium	2	5	8
	Low	4	7	9

Bibliography

Henderson, B. D. (1970) The product portfolio. The Boston Consulting Group, Perspective No. 66, Boston.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

cross-functional team

C. Anthony Di Benedetto

One of the most fundamental characteristics of new-product development is that the development is done by cross-functional teams. The activities that must be done for new-product success require skill in marketing, forecasting (*see* NEW-PRODUCT FORECASTING), research and development or engineering (*see* RESEARCH & DEVELOPMENT), production and manufacturing, operations, PRODUCT DESIGN, and these are among the functional areas within the firm that comprise the cross-functional team. Having a dedicated, empowered cross-functional team, and an effective team leader, is one way of accelerating the product's time to market (Gupta and Souder, 1998).

The teams, by definition, include individuals with different backgrounds and training, and this can cause trouble. Marketing people may fail to communicate effectively with R&D people, and vice versa (Souder, 1988). Different team members will have conflicting measures of success, or different expectations in terms of completion times. In some cases, individuals from different functional areas will be suspicious of or even dislike one another. Effective team management means managing the interfaces between the functional areas, and identifying and managing the frictions. Good managers know that a little conflict can be healthy, as it can lead to critical questions being raised internally and, possibly, some new solutions being identified. Nevertheless, it is important that conflicts among team members be properly managed. Forcing a solution, or seeking a superficial solution, is seen as less conducive to a positive environment for new-product development, while reaching an acceptable compromise or collaborative problem solving is preferable (Gobeli, Koenig, and Bechinger, 1998; Dyer and Song, 1998).

Cross-functional team management becomes substantially more complicated when the team has representatives from around the world. Firms are increasingly taking a global perspective to new-product development, tapping into knowledge and expertise that may reside in distant subsidiaries, and facing the challenges of managing global new-product teams. The barriers to overcome are many. Not all participants may be equally comfortable in the language used for team meetings; the meetings themselves are rarely in person but usually carried out electronically; and there are time zone differences to contend with. Cultural differences also exist, of course, which create more communication difficulties, but which may also lead to more creativity, multicultural synergies, and better problem solving (Song and Parry, 1997; Smith and Blanck, 2002; Sivakumar and Nakata, 2003).

Bibliography

Dyer, B. and Song, M. (1998) Innovation strategy and sanctioned conflict: a new edge in innovation? *Journal of Product Innovation Management*, **15** (6), 505–519.

Gobeli, D.H., Koenig, H.F., and Bechinger, I. (1998) Managing conflict in software development teams: a multi-level analysis. *Journal of Product Innovation Management*, **15** (5), 423–435.

Gupta, A.K. and Souder, W.E. (1998) Key drivers of reduced cycle time. *Research-Technology Management*, **41** (4) 38–43.

Sivakumar, K. and Nakata, C. (2003) Designing new global product teams: optimizing the effects of national culture on new product development. *International Marketing Review*, **20** (4), 397–445.

Smith, P.G. and Blanck, E.L. (2002) From experience: leading diverse teams. *Journal of Product Innovation Management*, **19** (4), 294–304.

Song, M. and Parry, M.E. (1997) Teamwork barriers in Japanese high-technology firms: the sociocultural differences between R&D and marketing managers. *Journal of Product Innovation Management*, **14** (5), 356–367.

Souder, W.E. (1988) *Managing New Product Innovations*, Lexington Books, Lexington, pp. 168–170.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

virtual teams

Richard Blackburn

Virtual Teams have been defined in as many ways as there are those who write about them. McShane and Von Glinow (2008) define *virtual teams* as "Teams whose members operate across space, time, and organizational boundaries, and who are linked through information and communication technologies (ICT) to achieve organizational goals" (p. 291). Virtual teams are also referred to as *distributed or geographically dispersed teams*, but whatever the terminology, such a collection of individuals is in contrast to collocated teams, all of whose members work in the same place and communicate face-to-face. The relatively widespread popularity of virtual teams comes in response to new demands in the organizational environment for the more flexible use of highly valued knowledge workers who must respond more quickly to competitive demands without incurring the costs and delays necessary to get distant individuals in one place at one time (*see also* ORGANIZING FOR INNOVATION).

An important distinction that should be made is between a virtual group and a virtual team. Virtual group members need not interact to complete their group task. Sales people who interact virtually with a sales manager can successfully accomplish their sales task without much interaction among group members. Virtual teams do require substantial levels of member interaction to be successful. A virtual new-product development team reflects a situation where substantive member interactions are important for team success. The effectiveness of virtual teams is influenced by many of the same factors as those that influence collocated teams. Thus, when putting such a team together, a manager must consider the team's task, size and composition, processes, and organizational culture. Concerns specific to virtual teams include type and availability of ICT and the levels of member and leader training in operating in the virtual environment (*see also* CROSS-FUNCTIONAL TEAM).

Tasks that are relatively straightforward and structured tend to be more appropriate for virtual teams than tasks that are amorphous and/or complex. While interdependence among team members is a given in virtual teams, attempting to manage substantial interactions with ICT is difficult. This also suggests that smaller teams composed of members with both cross-cultural sensitivity and skills to manage ICT will have a greater probability of being successful. These latter requirements often mean that certain training opportunities for would-be virtual team members and leaders will be beneficial. The best collocated team members may not make good virtual team members without specialized training.

An organization that wishes to organize itself around collaborative work units must insure that its systems and culture are supportive of these activities. Human resources systems like selection, training, compensation, and performance evaluation must be designed to enhance and reinforce the skills necessary to be a good virtual team member/leader. Additionally, the organizational values that undergird a firm's culture must support teamwork and multiple team membership for virtual teams to be successful.

Finally, processes in virtual teams differ somewhat from those found in collocated teams. All teams go through a developmental process beginning with a forming stage during which members meet and learn about each other. Occasionally, teams must work through a series of conflicts concerning the roles of various members and the procedures that the team will follow. This storming stage (hopefully) allows for the team to arrive at an agreement on how they intend to operate, and the norms that will guide these operations. Norm agreement and resolution of conflicts enables the team to begin performing at beneficial levels. In collocated teams, these stages tend to occur relatively quickly. Given the nature of the ICT, this process often can take more time with virtual teams. To move this developmental process along, virtual teams may meet together initially for a short time so that many of these issues can be resolved face-to-face. Once trust, norms, and cohesiveness are established with this early face-to-face encounter, it is much easier to maintain these processes in the virtual environment.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Bibliography

Blackburn, R., Furst, S., and Rosen, B. (2003) Building a winning virtual team. KSAs, selection, training, and evaluation, in *Virtual Teams that Work: Creating Conditions for Virtual Team Effectiveness* (eds C. Gibson, and S. Coben), Jossey-Bass, San Francisco, pp. 95–120.

Duarte, D. and Snyder, N. (2006) *Mastering Virtual Teams: Strategies, Tools, and Techniques that Succeed*, Jossey-Bass, San Francisco.

Furst, S., Reeves, M., Rosen, B., and Blackburn, R. (2004) Managing the life cycle of virtual teams. *Academy of Management Executive*, **18**, 6–20.

McDonough, E., Kahn, K., and Barczak, G. (2001) An investigation of the use of global, virtual, and collocated new product development teams. *Journal of Product Innovation Management*, **18**, 110–120.

McShane, S. and Von Glinow, M. (2008) *Organizational Behavior*, McGraw-Hill Irwin, Boston.

Montoya, M., Massey, A., Hung, C., and Crisp, C. (2009) Can you hear me now? Communication in virtual product development teams. *Journal of Product Innovation Management*, **26**, 139–155.

takeoff

Christophe Van den Bulte

Definition

In the realm of innovation and new-product marketing, takeoff is the first marked increase in sales. It corresponds to the transition between the introduction and growth stages of the PRODUCT LIFE CYCLE, and to the kink in the hockey-stick pattern of new-product or market sales. When the sales evolution of the product experiences multiple growth spurts, then the concept of takeoff is often limited to the first such spurt.

Relevance

Takeoff is of great importance to companies. Since sales and hence cash inflow dramatically increase after takeoff, knowing how long it will take for a new product to take off is very valuable to companies concerned about managing scarce financial resources. Knowing when takeoff will occur is also important for capacity planning in manufacturing, logistics, marketing communications (including sales force), and after-sales service. Given these cash-flow implications, knowing the likely time of takeoff is valuable to investors, too, and research has indeed documented abnormally high stock price returns in the year prior to takeoff (Markovitch and Golder, 2008). Having some consensus within the company of when a new product is likely to take off also gives business development and marketing managers the benefit of the doubt in the early stages of new-product commercialization, and prevents top management from "pulling the plug" and withdrawing the product (or market development investments) prematurely. Time to takeoff has also gained some popularity among marketing academics as a measure of diffusion speed, at least in the early stages of the new-product diffusion process (*see* DIFFUSION OF INNOVATION).

Measurement

The concept of takeoff has proved very hard to translate into a clear operational measure, and no consensus has emerged. Agarwal and Bayus (2002) and Golder and Tellis (1997), for instance, both measure takeoff as the first year in which a product's relative growth rate crosses a particular threshold value, but determine that threshold differently, with the former using a fixed value and the latter using a sliding scale rule. Tellis, Stremersch, and Yin (2003) also use a growth-rate threshold approach, but with a different sliding rule. Chandrasekaran and Tellis (2008) simply define takeoff as the year in which the market penetration reaches 2% of the population, whereas Garber *et al.* (2004) use 16% penetration as their criterion. Such inconsistencies across studies limit the ability to generate cumulative evidence and insights one can be confident about.

Findings

As new-product diffusion research has long documented, truly novel product categories often exhibit a pronounced S-shaped diffusion curve (*see* TECHNOLOGY S-CURVE; BASS MODEL), implying the presence of a takeoff taking place quite some time after the first product was introduced. Analyzing US sales data for 30 product categories, Agarwal and Bayus (2002) found that takeoff occurred on average 14 years after launch. They also found evidence that it tended to occur sooner over time, with the average time to takeoff decreasing from almost 19 years for products launched before World War II (WW II) to 10 years for those launched after WW II. Golder and Tellis (1997), using a different metric of takeoff, reported that time to takeoff for 31 consumer durables in the United States ranged from 2 to 60 years, with an average of 12 years (18 years before WW II and 6 after WW II). Analyzing the time to takeoff of 16 new consumer durables across 31 countries, Chandrasekaran and Tellis (2008) reported a steady and almost perfectly linear shortening of the time to takeoff from about 14 years for products introduced around 1915 to about four years for those introduced around 1995. Both these long periods to takeoff and the shortening over time are consistent with new-product diffusion research taking into account the entire diffusion process rather than only its earliest phase (Van den Bulte, 2000; Van den Bulte and Stremersch, 2004).

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Once one moves beyond average duration and acceleration and considers likely causes, consensus disappears. For instance, there is only weak and mixed evidence on whether takeoff occurs more quickly in countries that are wealthier. There is also disagreement about what specific features of national culture, if any, are associated with fast versus slow takeoff. Several reasons may operate for these weak and inconsistent findings: the lack of strong genuine effects, inconsistencies in how the studies were designed and takeoff was operationalized, and low statistical power. All three reasons may also explain why studies on new-product diffusion taking into account the entire diffusion process have led to more consistent findings (Talukdar, Sudhir, and Ainslie, 2002; Van den Bulte and Stremersch, 2004).

Of particular relevance to managers is the question how price drives time to takeoff. Conventional wisdom and the early study by Golder and Tellis (1997) suggest that price declines are likely to trigger takeoff, with every 1% decrease in price leading to a 4.2% increase in the probability of takeoff. More recent research, however, casts doubt on this explanation. Specifically, Agarwal and Bayus (2002) and Bayus, Kang, and Agarwal (2007) have presented results in which the effect of price on the time of takeoff vanishes once one controls for the entry of new competitors. This suggests that it need not be price itself but the presence of increased competition – which happens to trigger price erosion – that makes takeoff more likely to happen.

One unresolved question, then, is how and why firm entry has an effect on takeoff time if it is not through price erosion. As Agarwal and Bayus (2002) note, increased competition is likely to lead to significant improvement in product quality, and that may very well be the factor triggering takeoff. Other intermediate mechanisms may be the legitimation of the industry and product category (as when IBM entered the personal computer market in 1981) and the improvement of the distribution infrastructure since intermediaries are more likely to invest in supporting a new product or technology when more sources of supply become available. More research on how supply-side factors, especially those that can be influenced by managers, affect takeoff would be useful.

Acknowledgments

The text benefited from helpful suggestions by Professor Peter Golder, New York University, Professor Renana Peres, Hebrew University of Jerusalem, and Professor Qiaowei Shen, University of Pennsylvania.

Bibliography

Agarwal, R. and Bayus, B.L. (2002) The market evolution and sales takeoff of product innovations. *Management Science*, **48**, 1024–1041.

Bayus, B.L., Kang, W., and Agarwal, R. (2007) Creating growth in new markets: a simultaneous model of firm entry and price. *Journal of Product Innovation Management*, **24**, 139–155.

Chandrasekaran, D. and Tellis, G.J. (2008) Global takeoff of new products: culture, wealth, or vanishing differences? *Marketing Science*, **27**, 844–860.

Garber, T., Goldenberg, J., Libai, B., and Muller, E. (2004) From density to destiny: using spatial dimension of sales data for early prediction of new product success. *Marketing Science*, **23**, 419–428.

Golder, P.N. and Tellis, G.J. (1997) Will it ever fly? Modeling the takeoff of really new consumer durables. *Marketing Science*, **16**, 256–270.

Markovitch, D.G. and Golder, P.N. (2008) Using stock prices to predict market events: evidence on sales takeoff and long-term firm survival. *Marketing Science*, **27**, 717–729.

Talukdar, D., Sudhir, K., and Ainslie, A. (2002) Investigating new product diffusion across products and countries. *Marketing Science*, **21**, 97–114.

Tellis, G.J., Stremersch, S., and Yin, E. (2003) The international takeoff of new products: the role of economics, culture, and country innovativeness. *Marketing Science*, **22**, 188–208.

Van den Bulte, C. (2000) New product diffusion acceleration: measurement and analysis. *Marketing Science*, **19**, 366–380.

Van den Bulte, C. and Stremersch, S. (2004) Social contagion and income heterogeneity in new product diffusion: a meta-analytic test. *Marketing Science*, **23**, 530–544.

product positioning

David Reibstein

Introduction

Firms need to determine how they want their products positioned in the minds of their targeted customers. This means they should determine on what dimensions and where on these dimensions the product should be perceived relative to competition. The actual product position depends on how the customers actually view the firms' offerings on the dimensions of relevance to them. Perceptions of offerings are formed whether a company wants it or not and whether the firm is proactive in trying to shape the perceptions or not.

The Role of the Marketing Mix in Product Positioning

Advertising and other communications efforts are to first make prospective customers aware of the product/service and then secondarily to help form the customers' perceptions of the product. But, communications is not the only element used to form the positioning of the product offering.

For example, one of Toyota's brands is the Prius. It is targeted at the environmentally concerned consumer. In 2009, it is probably viewed as the most fuel-efficient automobile in the market, as it was one of the first hybrid cars introduced. It gained this positioning not simply through advertising. First, because it was the first hybrid introduced into the market, fuel efficiency was closely associated with the brand. So, time to market plays a role in positioning of a product. The firm *first to market* with a particular feature is often viewed most closely associated with that dimension (*see also* FIRST-MOVER ADVANTAGE).

Obviously, *the product,* or overall offering, plays a significant role in the resulting positioning. The fact that the Prius is a hybrid and gets good mileage per gallon of gasoline contributes significantly to the positioning of the brand as being fuel efficient. In this case, it is made easier, because there are objective measures, posted online and in the automobile dealers' showrooms, displaying the miles per gallon (mpg) of each model. More often than not, the critical dimensions are not objective and the task of positioning is more challenging. For example, Porsche wants to be positioned as sporty and Lexus luxurious and comfortable. Each of these has tried to achieve these positions and perceptions in customers' minds, by their product design (*see* PRODUCT DESIGN).

But, often, the product itself is not sufficient. Firms try to shape the customers' perceptions prior to the customer actually experiencing the product through their *communications* efforts. The Prius received considerable attention and notoriety around the fuel-efficient dimension because of the considerable public relations coverage they received. This has been coupled with an aggressive advertising campaign to support the continuing perceptions that the Prius is excellent in terms of fuel efficiency.

Similarly, the *distribution channel* with the showrooms, the signage at the showrooms, the sales people, and so on. all emphasize the fuel efficiency of the car. Where the product is distributed and how it is represented by the channel all contribute to the positioning.

The last element of the marketing mix is *price*. Prius charges a premium price. They justify this by the savings the customers will experience with the fuel efficiency. As such, even the price helps reinforce the perceptions in the marketplace.

The Role of Competition in Product Positioning

The Prius is no longer the most fuel-efficient car on the market. That said, the Prius still remains solidly in the *position* of being dominant in terms of fuel efficiency. The point is that while Prius competitors have technologically leapfrogged the Prius in terms of fuel efficiency, they have not in terms of positioning. The point is that positions are perceptual and these perceptions are not just shaped by the product alone. The critical point is that customers buy on the basis of their perceptions of the choices they have. The second point is that positions are also relative. What we consider as "long" battery life for a computer today is far different than what we did consider as a dominant position only a few years ago, and

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

what we will consider as such a decade from now. For a laptop to have a battery life of four hours in the mid-2000s was considered excellent, but today is on the low range. By the year 2015, it will not be surprising to find that average PC battery life will be well in excess of 24 hours and beyond. Hence, a firm's position is temporal and will be shaped by competitive activity.

General Motors will be offering their new Volt brand in 2010 which will have a fuel efficiency of 230 mpg, well in excess of the Prius' 40–51 mpg. With this introduction, without a significant change by Toyota, the Prius will no longer be perceived as the dominant fuel-efficient offering.

Selecting a Position

A firm needs to know where it should be positioned. This starts with (i) a selection of a target market, (ii) the identification of the critical dimensions the target uses for selecting which product it will choose, (iii) an understanding of the ideal levels the target desires on those dimensions, often referred to as an *ideal point*, and (iv) working through the use of the entire marketing mix to gain targeted customer perceptions closer to this ideal than any of the other choices on the market.

Bibliography

Kotler, P. and Keller, K. (2009) *Marketing Management*, Pearson-Prentice Hall, Upper Saddle River.

Ries, A. and Trout, J. (1981) *Positioning, The Battle for Your Mind*, Warner Books - McGraw-Hill Inc., New York.

open source

Nikolaus Franke

In *legal* terms, "open source" means that anyone who wishes can use, modify, and improve the product, and redistribute it in modified or unmodified versions to others. This term is most commonly used in the context of software to denote systems or applications (e.g., Linux, Apache, or Firefox) where the source code can be accessed by any user, and which are distributed under specific copyright licenses that ensure the above-mentioned privileges (for a definition, see http://opensource.org/docs/osd). However, the principle of not asserting private intellectual property claims (*see* INTELLECTUAL PROPERTY RIGHTS) can be applied to other products as well.

In *economic* terms, open-source software is a public good. Because of its modular character (*see* PRODUCT MODULARITY), it is often developed by virtual-user networks or communities in a distributed and collaborative manner. Users are not paid in exchange for their contributions, and the software can be downloaded freely by anyone. Prior to the success of open-source software, many held a strong belief that such a regime of systematic "free-riding" would bring about insufficient incentives to innovate. However, in many cases, open-source software is highly innovative and successful, and its importance is still growing. More than 100 000 open-source projects are currently underway, and estimates of the value created by such projects run into billions of dollars. Open-source software can be considered one of the strongest arguments for the importance of the mega trend toward OPEN INNOVATION. Its success has been explained by the fact that developers themselves decide which problem they work – the assignment thus is not made by a central authority. Often, the specific problem is far ahead of current trends, and the user urgently needs a solution himself/herself – thus many contributions can be referred to as lead-user innovations (*see* LEAD USERS). Second, a potentially large number of users contribute to open-source projects – far more than in proprietary software production. The advantage is summarized in what Raymond (1999) termed *Linus' Law*: "Given enough eyeballs, all bugs are shallow."

All this raises the question of user *motives* for contributing to this public good. Research has found that both intrinsic (e.g., fun, enjoyment, altruism) and extrinsic motives (e.g., learning, reputation building, social norms) explain why users participate and freely reveal their proprietary knowledge.

A great deal of research has been devoted to the question of how open-source projects are organized and governed. Compared to hierarchical systems such as companies, there is a relatively low level of coordination activity, and the most important management tool is a simple mailing list. Contribution activities are often highly concentrated among a small number of core people, but the projects involve other important roles as well (such as testing and assistance). This loose form of governance sometimes leads to a "forking" of the project, that is, two (or more) versions of the software are developed in parallel.

For companies, the open-source phenomenon can present a challenge as well as a business opportunity. First, the provision of a free, high-quality public good threatens the existence of companies that operate by producing and selling similar (proprietary) products. On the other hand, there are many potential business models that would allow companies to profit from open-source software (and other "open-source" products). One possible model would be to use the software for one's own business purposes, while another would be to market complementary products and services. Companies can also adopt the principles of open-source projects and outsource certain elements of their own new-product development and marketing activities to self-organized (or loosely governed) communities of users. Such activities might include idea, concept and prototype generation, RESEARCH & DEVELOPMENT, CONCEPT TESTING, and PRODUCT TESTING, customer support and help-line activities, and so forth.

Bibliography

Bonaccorsi, A. and Rossi, C. (2003) Why open source software can succeed. *Research Policy*, **32**, 1243–1258.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Fleisher, C.S. (2008) Using open source data in developing competitive and marketing intelligence. *European Journal of Marketing*, **42**, 852–866.

Franke, N. and von Hippel, E. (2003) Satisfying heterogeneous user needs via innovation toolkits: the case of apache security software. *Research Policy*, **32**, 1199–1215.

Hertel, G., Niedner, S., and Herrmann, S. (2003) Motivation of software developers in open source projects: an internet-based survey of contributors to the Linux kernel. *Research Policy*, **32**, 1159–1177.

von Hippel, E. and von Krogh, G. (2003) Open source software and the "private-collective" innovation model: issues for organization science. *Organization Science*, **14**, 209–223.

von Krogh, G. and von Hippel, E. (2006) The promise of research on open source software. *Management Science*, **52**, 975–983.

von Krogh, G., Spaeth, S., and Lakhani, K.R. (2003) Community, joining, and specialization in open source software innovation: a case study. *Research Policy*, **32**, 1217–1231.

Lakhani, K.R. and von Hippel, E. (2003) How open source software works: "free" user-to-user assistance. *Research Policy*, **32**, 923–943.

Lerner, Josh. and Tirole, Jean. (2002) Some simple economics of open source. *Journal of Industrial Economics*, **50**, 197–234.

Pitt, L.F., Watson, R.T., Berthon, P. *et al*. (2006) The penguin's window: corporate brands from an open-source perspective. *Journal of the Academy of Marketing Science*, **34**, 115–127.

Raymond, E. (1999) *The Cathedral and the Bazaar*, O'Reilly, Sebastopol.

technology S-curve

Ashish Sood

The Technological S-curve

The technological S-curve describes the phenomenon of technological evolution and suggests that technologies evolve through an initial period of slow growth, followed by one of fast growth culminating in a plateau (Foster, 1986; Sahal, 1981; Utterback, 1994). When plotted against time, the performance resembles an S-curve (see Figure 1).

The field does not enjoy a single, strong, and unified theory of technological evolution. However, an emerging, and probably the most compelling explanation, revolves around the dynamics of firms and researchers as the technology evolves through the three major stages of the S-curve of technological evolution: introduction, growth, and maturity (*see also* PRODUCT LIFE CYCLE).

Introduction Stage

A new technological platform (Sood and Tellis, 2005) initially makes slow progress in performance during this early phase of its product life cycle. Two reasons may account for this situation. First, the technology is not well known and may not attract the attention of researchers. Second, certain basic but important bottlenecks need to be overcome before any new technological platform can be translated into practical and meaningful improvements in product performance.

Growth Stage

With continued research, the technological platform crosses a threshold after which it makes rapid progress. Three factors may account for this change—the emergence of a dominant standard (Utterback, 1974); product characteristics and consumer preferences coalesce on the new standard and/or larger number of researchers attracted by the publicity of the standardization; and increase in sales of products that translates into greater support for research.

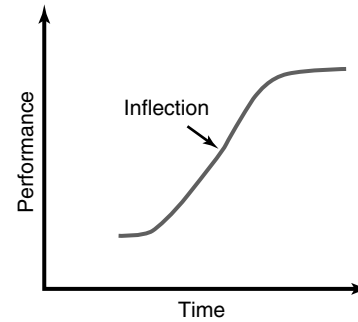


Figure 1 Technological evolution: technological S-curve.

Maturity Stage

After a period of rapid improvement in performance, prior research suggests that the new technology reaches a period of maturity when progress occurs very slowly or reaches a ceiling (Foster, 1986; Brown, 1992; Utterback, 1994; Chandy and Tellis, 2000) for various reasons—innate characteristics of a technology; changing focus of innovation as markets saturate; fears of obsolescence or cannibalization; and limits of scale or system complexity.

Recent Empirical Findings

Recent empirical findings on the phenomenon question the support for the theory (Sood and Tellis, 2004, 2005; James and Sood, 2005). These authors plotted the performance of 23 technologies over time from six markets over more than 100 years: four each in desktop printing, display monitors and analgesics, and three each in desktop memory and data transfer, and five in external lighting. In a majority of the technologies, they found long periods of static performance interspersed with abrupt jumps in performance. This suggests that the path of technological evolution is neither smooth nor does it resemble an S-curve.

If a technology does not evolve in the shape of an S-curve, an analyst expecting such a path might conclude from periods of static performance that the technology has matured at the upper asymptote, when indeed it has not. The risk is in abandoning the old technology prematurely.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Bibliography

Brown, R. (1992) Managing the 'S' curves of innovation. *Journal of Consumer Marketing*, **9** (1), 61–73.

Chandy, R.K. and Tellis, G.J. (2000) The incumbent's curse? Incumbency, size, and radical product innovation. *Journal of Marketing*, **64** (3), 1–17.

Foster, R. (1986) *Innovation: The Attacker's Advantage*, Summit Books, New York.

James, G. and Sood, A. (2005) Performing hypothesis tests on the shape of functional data. *Computational Statistics and Data Analysis*, **50** (1), 1774–1792.

Sahal, D. (1981) Alternative conceptions of technology. *Research Policy*, **10** (1), 2–24.

Sood, A. and Tellis, G.J. (2004) The S-curve of Technological Evolution: Strategic Law or Self-Fulfilling Prophecy? MSI Reports #4-1223, Marketing Science Institute, Cambridge.

Sood, A. and Tellis, G.J. (2005) Technological evolution and radical innovations. *Journal of Marketing*, **69** (3), 152–168.

Utterback, J.M. (1974) Mastering the dynamics of innovation. *Science*, New Series, **183** (4125), 620–626.

Utterback, J.M. (1994) *Mastering the Dynamics of Innovation*, Harvard Business School Press, Boston.

product modularity

Tucker J. Marion

Definitions

Modularization is an approach to organize complex designs and process operations more efficiently by decomposing complex systems into simpler subsystems (Jose and Tollenaere, 2005). These subsystems may be a common engine–transaxle combination used in a variety of automobiles. For example, Honda shares an engine and transaxle combination among a variety of vehicle lines including the compact Civic, midsize Accord, and CR-V sport utility (Meyer, 2007). What makes this possible is well-defined product architecture and component interfaces (*see also* PRODUCT DESIGN).

Product architecture is the method by which product function ties to physical components. Defining product architecture is a key element in the engineering design of modular products, and should be performed during system-level design early in the new-product development process (Ulrich and Eppinger, 2004). Ulrich (1995) defines product architecture as (i) the arrangement of *functional elements*; (ii) the mapping from *functional elements* to *physical components;* and (iii) the specification of the *interfaces* among interacting physical components. The functional elements are high-level product features and actions. For example, a functional element in a door handle is to *secure door*. The corresponding physical component that will secure the door is a *latch*. The latch will need to have well-defined connection points, or interfaces. In this example, thought and specification will need to go into the methods by which the latch connects to the door handle, locking mechanism, and door jamb.

To accomplish the arrangement of functional elements into building blocks and define their interfaces, several methods have been proposed, including the development of functional diagrams (Pahl and Beitz, 1996). The function diagram or function structure is a representation of an electromechanical system consisting of inputs to the system, outputs from the system, and internal subfunctions within the system, which can then be used to identify pertinent interfaces within the system (Pahl and Beitz, 1996). Physical components tie to functions and other components in the system. A functional diagram of the door handle example is shown in Figure 1 (Marion *et al.*, 2006).

As shown in Figure 1, interfaces are interacting components (i.e., a door handle/knob connected to a latch) with geometric connections. An interface specification defines the protocol for the primary interactions across the component interfaces (Ulrich, 1995). For example, if one were designing a door handle and locking system, they might want to offer the option of different door handles/knobs to the consumer. This requires a common interface between the components. Figure 2 illustrates a common door handle/knob and shaft interface.

Standard interfaces are common in industry; examples include USB connections and automobile tire hubs. A product is termed *modular* if there is a one-to-one or many-to-one mapping of functional elements to physical components (Ulrich, 1995). In the door handle example, to actuate the system, one must turn the door handle/knob. The latch performs the locking function. This is an example of a modular design via one-to-one mapping of function to component. Modular subsystems and design are applicable to other areas such as software development. For example, the Mach kernel in Apple's OS X operating system is a software subsystem that is ported to the iPhone, iPod Touch, and Mac personal-computer platforms (Silbershatz *et al.*, 2004).

Benefits

Because modular designs have a beneficial ratio of functions to components (one-to-one in the extreme), components of the product can be modified or replaced without changing other components in the design (Ulrich, 1995). For example, engines can be upgraded over time in the Honda example. This can focus R&D resources on localized components or subsystems that need change, thereby reducing the need to modify the entire product. Modular products have well-defined interfaces, so they have the inherent property that they can be manufactured in a flexible manner. Again using the Honda case, vehicle assembly plants can produce a wide variety of different models with

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

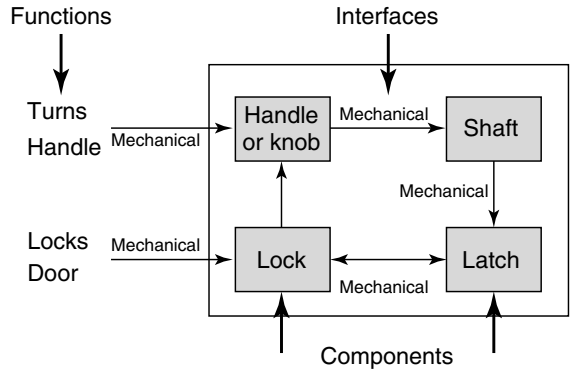


Figure 1 Example of a functional diagram of a door handle/knob noting the functions, components, and interfaces.

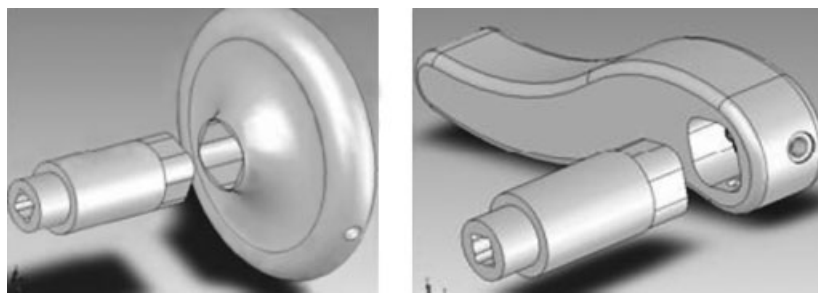


Figure 2 Interface design of a door handle/knob and shaft; note the common interface and unique handle/knob.

modularity, reducing changeover cost and maximizing economies of scale (*see also* PRODUCT-LINE STRATEGIES).

Product modularity is closely related to another engineering paradigm: the PRODUCT PLATFORMS. Product platforms involve the use of common components and subsystems throughout a product line (Meyer and Lehnerd, 1997). Product modularity at the component and system levels can foster subsystem and component reuse.

In summation, designing a modular product has many benefits from manufacturing flexibility to reduced R&D expenses during PRODUCT DESIGN. Well-defined product architecture and interfaces can foster the development of product platforms, which can help a firm easily introduce successive revenue-generating variants. During up-front product design, intelligently designing modularity into components and subsystems is an essential step toward development of a successful product line.

Bibliography

Jose, A. and Tollenaere, M. (2005) Modular and Platform methods for product family design: literature analysis. *Journal of Intelligent Manufacturing*, **16** (3), 371–390.

Marion, T.J., Freyer, M., Simpson, T.W., and Wysk, R.A. (2006) Design for mass customization in the early stages of product development. 2006 DETC, ASME Conference Proceedings, Philadelphia, PA, DFMLC-99641.

Meyer, M.H. (2007) *The Fast Path to Corporate Growth: Leveraging Knowledge and Technologies to New Market Applications*, Oxford University Press, New York.

Meyer, M.H. and Lehnerd, A.P. (1997) *The Power of Product Platforms: Building Value and Cost Leadership*, Free Press, New York.

Pahl, G. and Beitz, W. (1996) *Engineering Design: A Systematic Approach*, 2nd edn, Springer-Verlag, New York.

Silbershatz, A., Galvin, P. B. and Gagne, G. *et al*. (2004) *The Mach System, Appendix to Operating Systems*, 7th edn, John Wiley & Sons, Inc., New York, ISBN 0471694665.

Silbershatz, A., Galvin, P. B. and Gagne, G. *et al*. http://www-2.cs.cmu.edu/afs/cs/project/mach/public/www/mach.html (accessed 28 October 2008).

Ulrich, K. (1995) The role of product architecture in the manufacturing firm. *Research Policy*, **24**, 419–440.

Ulrich, K.T. and Eppinger, S.D. (2004) *Product Design and Development*, 3rd edn, Irwin McGraw-Hill, New York.

prototype

Richard Blackburn

Prototype (*prototyping*) refers to an approximation or model (in any appropriate medium) of an idea or concept along one or more dimensions of interest. Prototypes can be produced at any point in any development process. Michael Schrage (2000b) argues that prototypes can serve as valuable "shared spaces" within which a new-product development (NPD) team might enhance the efforts of its development process. While concept specifications might describe the team's "wish list" (*see* PRODUCT SPECIFICATIONS), prototypes make those wishes "real." Prototypes might suggest where the team's wishes were unrealistic, but they can also indicate where the team's wishes might not have been imaginative enough (Schrage, 2000b, p. 71).

Prototypes come in a variety of flavors – as simple as a diagram on the back of a napkin or as complex as a fully functional representation of a product one step removed from actual production.

More specifically, NPD teams can produce two-dimensional drawings or computer-generated plans, three-dimensional "form studies" (built from foam, polymer, or constructed "virtually" with computer-assisted design systems), functional prototypes showing how a product is supposed to work but without any "packaging," user test models that incorporate function and packaging, and system prototypes that illustrate how manufacturing, packaging, and distribution systems must be aligned to incorporate the new product (*see also* PRODUCT TESTING).

As the team moves from simple to more-complex prototypes, the cost of doing so can increase substantially. However, recent advances in rapid prototyping provide less-expensive methods for building more-complex models.

Judicial use of prototypes can assist the NPD team in a number of ways. For instance, *prototypes can help the team learn* if the proposed concept works as anticipated, and/or if the concept meets the customer's needs or wants. *Prototypes can facilitate communication* among NPD team members and among other key stakeholders both inside and outside the team. A discussion around the tangible rendering of any concept will likely be more useful than a discussion in which ideas remain as evanescent dreams in team members' minds. *Prototypes can assist in the integration of information* about the fit of components in a larger concept as well as the extent to which cross-functional demands are being met in a particular concept. Finally, *prototypes can function as milestones*, indicating how far the NPD team has progressed toward a development goal and/or the extent to which the team has proven the value/worth of the proposed concept.

Interestingly, Schrage suggests that just as organizations may have quality or safety cultures, they can also have a "prototyping culture," which he contrasts with a "specification culture." Firms with specification cultures often tend to innovate at the margins of existing products. Those with prototyping cultures are more likely to produce substantive innovations with unique offerings. From Schrage's perspective, strong prototyping cultures produce truly innovative products (*see* RADICAL INNOVATION; INNOVATION TYPOLOGIES).

If high-quality prototypes are a function of a company's culture and its willingness and ability to innovate, what are some of the characteristics of such cultures? In prototyping cultures (versus specification cultures), the prototypical models are

- referred to as *prototypes* and are *designed* rather than being called *mock-ups* that are *assembled*;
- used to explore opportunities and provide feedback rather than manage risks and publicize products;
- ephemeral and dynamic (virtual) rather than permanent and static (actual);
- appear early and often at various steps in the NPD process rather than infrequently and late in that process;
- not "frozen" until late in the process rather than frozen whenever they appear in the process;
- community property open to public viewing at any time rather than being "owned" by

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

a particular function/unit with controlled viewing; and

- developed with frequent external contribution/collaboration rather than developed solely with contributions from the NPD team (Schrage, 1993).

In sum, if the concept of quality is approximately a function of the quality of prototypes produced multiplied by the number of prototypes produced, then to build better concepts NPD teams must learn to build better prototypes and to build them faster.

Bibliography

Roam, D. (2008) *The Back of the Napkin: Solving Problems and Selling Ideas with Pictures*, Penguin Group, New York.

Schrage, M. (1993) The culture(s) of prototyping. *Design Management Journal*, **4** (Winter), 55–65.

Schrage, M. (2000a) Serious play: the future of prototyping and prototyping the future. *Design Management Journal*, **11** (Summer), 29–37.

Schrage, M. (2000b) *Serious Play: How the World's Best Companies Simulate to Innovate*, Harvard Business School Press, Cambridge.

Ulrich, K. and Eppinger, S. (2007) Prototyping, in *Product Design and Development*, 4th edn, Chapter 12, McGraw-Hill/Irwin, New York.

product specifications

Robert W. Veryzer

This entry encompasses Product Specifications as the culmination of an evolution from determining product Requirements to the development of Objectives that provide guidance for detailed descriptions for how a product is to be constructed. Although terminology varies somewhat across industries and firms, with "product requirements," "engineering characteristics," "specifications," or "technical specifications" being used in some instances (Ulrich and Eppinger, 2008), the steps that result in product specifications are essentially the same. In effect, "product specifications" are the outcome of a process that leads to a precise blueprint for (making) a particular product (*see* PRODUCT DESIGN).

Requirements (e.g., "easy to operate," "large storage capacity," "light weight") are statements of what has been determined through study and exploration of user/customer needs (*see* VOICE OF THE CUSTOMER), and relevant market or firm considerations for a new product (Veryzer, 2003). These may come from knowledge of the marketplace, interviews with LEAD USERS or existing or potential customers, input from distribution channel members (e.g., dealers or retailers), research or product use tests, and so on. Requirements may also encompass such aspects of a product being designed as the physical product structure, operating systems, interactions among or dependencies with other products, and others, to the extent that these are indicated through investigation of one or more of the input sources (e.g., lead users, dealer network, and product-use testparticipants).

After a development team (*see* CROSS-FUNCTIONAL TEAM) has devoted sufficient effort to identifying and understanding user/customer needs (*see* VALUE PROPOSITION), and then identified the *requirements* for a product being designed, the requirements are formulated into descriptions of particular aspects of the product. Although various terms may be used, these descriptions are essentially *objectives* that indicate the formal goals for various aspects of a product such as performance capabilities and task times. Either formally or implicitly, the determination of objectives involves judgments concerning trade-offs among different product characteristics (e.g., speed vs. weight).

Objectives, in turn, are translated into precise targets, or *product specifications*. Product specifications detail in precise, measurable ways how a product is to be made. Product specifications – which may range over various dimensions of a product: physical elements and materials, structure or architecture, control or operating systems, and so on – consist of metrics or measures relevant to a particular product dimension, and indicate the desired level (or value) that is to be achieved for each measure. In the translation of objectives to specifications, the capturing (measuring) of more subjective needs, though challenging, should be done in as concrete a way as possible (for example, an index or assessment of customer satisfaction).

The process (*see* THE STAGE-GATE IDEA TO LAUNCH SYSTEM) for determining product specifications may vary depending on the type of product being designed (for example, consumer durables vs. electronics products vs. computer software). However, the following considerations usually apply:

- The general sequence for determining product specifications moves from requirements to objectives to product specifications.
- Each class or type of product may have its own standards and conventions.
- Marketing research can play an important role in the identification and validation of requirements.
- Objectives should recognize resource and time constraints.
- For certain types of products (for example, software and electronics), there may be established, very particular formats for requirements and product specifications.
- With advances in electronic and digital technologies, there is now often considerable latitude in specifications as concerns product configuration (in general, electronics products and software tend to be less constrained than products based on mechanics).

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Bibliography

Ulrich, K.T. and Eppinger, S.D. (2008) *Product Design and Development*, The McGraw-Hill Companies, Inc., New York.

Veryzer, R.W. (2003) Marketing and the development of innovative new products, in *The International Handbook on Innovation* (ed. L.V.Shavinina), Elsevier Science Ltd, pp. 845–855.

process innovation

Henry Robben

Process innovations typically change the way in which suppliers create or distribute products or services (Tushman and Nadler, 1986). In essence, the product or service remains the same, in terms of its appearance or the benefits it provides. It is the way of their production that changes. Typically, most of such changes are what we call *incremental* or *synthetic* changes. They lead to an improvement in production costs or product quality. As such, they improve the efficiency of current production processes. We also discern radical or discontinuous process innovations (*see* RADICAL INNOVATION; INNOVATION TYPOLOGIES). Such innovations significantly alter the way in which we produce products or services. For those innovations to occur, we need different approaches in management and manufacturing, and, likely, new skills sets as well.

Process innovation is especially important during the growth phase of the product life cycle, when competition revolves around price, quality, and segmentation (*see* PRODUCT LIFE CYCLE). A consequence of this may be the opening up of a market for a larger audience and product enhancement. As such, process innovation may play a strategic role for the organization.

The performance of a new-product development project tends to be determined by a few high-level, critical processes that need management's focused attention and efforts. To make it work well, all participants should understand the whole process and be motivated; otherwise; it cannot be effectively executed or improved. It is necessary to have a thorough understanding of the business and of the company's innovation management process to be able to improve, reengineer, or innovate any process within the organization.

The task of all managers is to improve their operations. There are several approaches to trigger off process innovation or improvement. We can distinguish the following triggers (Trott, 2008):

1. *Gap analysis*: Are there differences in expectations among customers, suppliers, and management?
2. *Quality circles and process improvement teams*: On the basis of their individual expertise, members of such groups suggest ways to innovate or improve current processes. They may work on a voluntary basis or be a systematic part of the innovation and new-product development processes.
3. *Total quality management (TQM)*: This is a system that continuously integrates quality development, quality maintenance, and quality improvement. It is an integral approach, requiring efforts from multiple departments, aimed at creating customer satisfaction at the most economical level.
4. *Quality function deployment (QFD)*: QFD is a concurrent methodology that systematically relates customer requirements to marketing, engineering, and operations requirements (*see* QUALITY FUNCTION DEPLOYMENT (QFD)).
5. *The ISO 9000 approach*: This is an internationally used system that checks whether an organization adheres to specific requirements.
6. *The European Foundation for Quality Management (EFQM) excellence model*: According to this system, a number of enablers lead to performance through managing and controlling input and output processes. It is a self-assessed procedure.
7. *Six sigma*: This is a methodology for designing complex systems, using an approach that defines, measure, analyzes, improves, and controls the extent to which results meet requirements. It should lead to a reduction of manufacturing or service problems in the parts-per-million range.

Process innovation measures and modifies organizational dynamics by optimizing the activities required to reach end goals or objectives (Tighe and Kraemer, 1996). A procedure to arrive at process innovation may be the following:

1. *Process mapping*: This phase shows how a process proceeds in time and the agents involved in the process. It delivers a flow

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

of how the agents influence and direct the process, and with what frequency.

2. *The ist-situation*: It is important to understand how the process currently works, and why. We need to evaluate what drives and maintains the current process.
3. *Benchmarking, best practices, next practices*: Research should show which companies or business units have an exemplary performance, either within or outside the industry. Here, we should emphasize that reaching a benchmark may not be sufficient in creating radical process innovation. Rather, we should look for next practices.
4. *The soll-situation*: Here, we need to have a good understanding of the desired situation. Why do we want or need to be there? This objective should be consistent with or leading to the company's long-term objectives. Sometimes, we call this process *blueprinting*.
5. *Implementation*: Obviously, a process innovation needs implementation. Here, management and workers face continuous changes, pilot testing and working with early adopters, training them and the workers, and documenting the process.

Bibliography

Tighe, G.S. and Kraemer, B.P. (1996) Using a concurrent team to reengineer the product development process, in *The PDMA Handbook of New Product Development* (ed. M.D. Rosenau Jr.), John Wiley & Sons, Inc., pp. 441–454.

Trott, P. (2008) *Innovation Management and New Product Development*, 4th edn. Prentice Hall Financial Times.

Tushman, M. and Nadler, D. (1986) Organizing for innovation. *California Management Review*, **28** (3), 74–92.

product diversification

Ashish Sood

Product diversification is one of the four main marketing strategies defined by the Ansoff matrix, a tool that helps firms formulate their product and market growth strategy (Ansoff, 1957) (see Table 1) (*see also* GROWTH STRATEGIES).

In this growth strategy, a firm markets new products in new markets. Diversification is the most risky of the four growth strategies since it requires a firm to acquire new skills, new techniques, and new facilities. In contrast, the other three strategies require existing competencies of the firm. As a result, diversification often leads to substantial changes in the organizational structure, competencies, and resources of the firm (*see also* ORGANIZING FOR INNOVATION; CORE COMPETENCIES). Therefore, a firm should choose this option only after careful analyses of all opportunities.

Types of Diversification

There are three types of product diversification: vertical, horizontal, and lateral (Ansoff, 1957). In vertical diversification, a firm introduces new products, which are related to current products, into new markets. Hence, the focus is on creating a portfolio of related businesses. In horizontal diversification, a firm develops or acquires new products that are different from its current products, but which may appeal to its current customers. Hence the focus is on gaining market share. In lateral diversification, a firm introduces new products that are different from its current products into new markets. Hence the focus is on entering completely new markets with new products and is more radical than the first two types of diversification.

Firms may employ these three types of diversification on the basis of market growth conditions (Ramanujam and Varadarajan, 1989). For example, a firm may choose either vertical or horizontal diversification when market conditions allow firms fast growth opportunities. On the other hand, lateral diversification may become necessary when market conditions show declining overall demand. Similarly, a firm may choose lateral diversification to develop competencies in a new area of technology.

Table 1 Ansoff Matrix.

		Products	
		Present	*New*
Markets	**Present**	Market penetration	Product development
	New	Market development	Product diversification

Bibliography

Ansoff, I. (1957) Strategies for diversification. *Harvard Business Review*, **35** (5), 113–124.

Ramanujam, V. and Varadarajan, P. (1989) Research on corporate diversification: a synthesis. *Strategic Management Journal*, **10** (6), 523–551.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

prediction markets

Thomas S. Gruca

Introduction

A prediction market is a market-based method for forecasting (*see* CONCEPT TESTING, NEW-PRODUCT FORECASTING, PRETEST MARKET MODELS). Prediction markets are used to generate a forecast regarding a specific future outcome such as the vote share in a political election, the opening weekend box office results for a new movie, or the winner of a contest (e.g., a World Cup final). Participants interact with each other by buying and selling financial instruments (i.e., futures contracts) whose ultimate value is determined by the outcome of the future event. The prices in a prediction market reveal the consensus forecast of the traders (Gruca, Berg, and Cipriano, 2005). In many applications, these forecasts have been highly accurate on an absolute basis and when compared to alternative forecasting methods. For example, the Iowa Electronic Market (IEM) has been forecasting the share of the popular vote for every presidential election since 1988. The mean absolute error of the forecasts for the 1988–2004 elections were 1.33% compared to 1.62% for the average public opinion poll. In 2008, the IEM was even better with a mean absolute error of 0.3% versus 1.2% for polls.

Over the past few years, the applications of market-based tools for forecasting have expanded in several important directions including new types of traders and different outcomes of interest. After reviewing the basic theory behind using markets as an information-gathering tool, the major advances in this area are discussed in approximate order of development.

Theoretical Background

Prediction markets take advantage of a market's ability to distill disparate sources of public and private information into a single measure: price. Rational expectations theory suggests that price alone is a sufficient statistic for all information available to traders in a market. Consequently, through the mechanism of price, all traders in a market (and any observers) can share the knowledge held by each and every trader (Sunder, 1995). In theory, markets should be an excellent method for aggregating information about future events and providing an accurate forecast.

Laboratory studies in experimental economics document the ability of markets (i) to disseminate information from informed traders to uninformed traders, and (ii) to aggregate information across traders. These laboratory results suggest that markets are capable of the information-processing tasks required by rational expectations theory (Sunder, 1995). The final question to be addressed is whether actual prediction markets could realize their potential to provide accurate forecasts of future events.

Initial (and Continuing) Success

The first public prediction market was the Iowa Political Stock Market (now known as the Iowa Electronic Market or IEM) conducted to predict the vote shares of the two major candidates in the 1988 US presidential election. The market's election eve predictions of the popular vote shares were more accurate than that those available from pre-election "trial heat" polls (Forsythe *et al.*, 1992).

One key finding from this market is that traders need not be a representative sample of voters in order for the market to provide accurate forecasts. In the 1988 market, the traders were primarily students and faculty at the University of Iowa. Even though these traders were clearly not a representative sample of voters, the prices in the markets predicted the election-day vote shares better than polls based on representative sample of voters.

A second key finding is that there is clear evidence that the average trader is biased (e.g., paying more for contracts associated with one's favored candidate) and prone to trading errors (e.g., creating risk-free arbitrage opportunities for other traders). Nevertheless, the IEM election markets have forecast the eventual outcomes (i.e., popular vote shares) with a higher degree of accuracy than the polls.

To explain how a market made up of biased, error-prone traders can produce accurate forecasts, Forsythe *et al.* (1992) proposed the "marginal trader hypothesis." Marginal traders

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

are able to put aside their personal preferences and trade according to their heads, not their hearts. They are usually very active traders who tend to be the ones setting prices in the market and they are less likely to make trading errors in favor of other participants. It is the price-setting actions of the marginal traders that move the prices of the contracts in prediction markets to correct levels.

Internal Prediction Markets

A major advance in the application of prediction markets was to apply this method to the problem of forecasting the future sales of existing products. These markets, conducted by researchers from the California Institute of Technology and HP Laboratories, used selected managers from Hewlett-Packard as traders (Chen and Plott, 2001). The idea was to test whether a prediction market could successfully aggregate important business intelligence that is widely scattered throughout a large organization. The forecasts generated by contract prices in the prediction markets were generally more accurate than those produced by the firm's own official forecasts.

These results showed that prediction markets are useful in situations where the traders were relying mainly on private information. In an election prediction market conducted in the United States, there is often a great deal of public information (e.g., trial heat polls) upon which to base one's trades. In contrast, the traders in the HP markets were bringing private information to the market from many parts of the overall organization. Through trading, these varied sources of information were aggregated into contract prices that yielded a consensus forecast as well as a reflection of the degree of differences in opinions across the traders.

Security Design in Real Money Prediction Markets

The value of the securities being traded in a prediction market ultimately depends on the outcome of interest. How the outcome is related to prices in the prediction market is a key design choice. For example, consider a prediction market focused on the popular vote in a US presidential election. Such a market may involve trading in three securities, one for each of the Democratic and Republican nominees and one for "All other candidates." One approach is to link the ultimate value of the security to the proportion of the total vote share for the corresponding candidate. At the end of the market, the total value of all three securities would be unity ($1).

A second alternative is a "winner-take-all" payoff structure. In this design, the security associated with the candidate that obtains the largest share of the popular vote pays off $1 and the other two securities payoff $0. Using this latter payoff structure is also known as offering a set of Arrow–Debreu securities. This approach has been used to provide forecasts of continuous outcomes such as product sales (Chen and Plott, 2001) and movie box office receipts (Gruca, Berg, and Cipriano, 2005).

To provide traders with contracts to sell, many prediction markets follow the design of the IEM by using a "bundle" transaction. For example, consider a prediction market with three mutually exclusive and collectively exhaustive outcomes. The bundle contains one contract corresponding to each of the possible outcomes of the future event. Traders buy a bundle of contracts from the market at a fixed price. This price is the same value as holding a "winning" contract, usually $1. Since one of the contracts will pay off while the others will become worthless, the expected value of a set of all contracts is $1. This approach ensures all transaction prices are set by traders and that the supply of contracts is limited only by the funds available to the traders.

Play Money Prediction Markets

A prediction market uses a market mechanism to aggregate the information available to individual traders. Traders with superior information about the outcome of interest can benefit from trading with other participants who are less well informed. The key question is how to motivate traders with valuable information to participate in a prediction market.

One possible solution is to use real money incentives. However, depending on the application, a public prediction market using real money payoffs may be subject to laws regarding gambling, securities trading, sales promotions,

and election fraud prevention, among others. Such laws vary widely across jurisdictions (Hahn and Tetlock, 2006).

The Hollywood Stock Exchange (HSX) is an internet-based game. The goal of the game is to acquire fictitious "Hollywood dollars" (denoted H$) through the trading of various securities tied to outcomes in the entertainment field including box office performance, awards shows, and so on. The most significant of these securities is the "Moviestock." The value of a Moviestock is determined by the total US box office receipts (in millions of dollars) over the first four weekends after a film has achieved wide release (greater than 650 screens). For example, if traders (as a group) expect a movie to take in $75 million during that time period, the price for the Movie stock will be H$75.

A Moviestock is usually issued very early in the development of a movie and its value changes as the film moves from concept to finished product. The initial price of a Moviestock is set by the HSX exchange. To avoid the influence of traders with inside information, trading is halted on Friday afternoon of the first weekend the movie is widely released in theaters. The Moviestock price is then reset by HSX to 2.9 times the first weekend's box office receipts (in millions of dollars) on Monday when trading resumes (different multipliers are used for movies opening on different days). Such market interventions are possible since HSX uses play money.

Empirical studies of the accuracy of the predictions of HSX markets find a surprisingly high level of forecasting accuracy (Pennock *et al.*, 2000; Spann and Skiera, 2003) on an absolute basis and when compared to the predictions of individual forecasters. A recent study comparing the results of 19 real money (IEM) and play money (HSX) box office prediction markets found no significant differences in overall accuracy (Gruca, Berg, and Cipriano, 2008).

Traders in play money prediction markets such as HSX are motivated by considerations other than monetary gain. By using play money rather than real money payoffs, users of prediction markets may be able to avoid the legal challenges noted above. However, it should be noted that some studies comparing the relative accuracy of real money and play money prediction markets find significant improvements when real money payoffs are used (Rosenbloom and Notz, 2005).

Introduction of User-friendly Trading

Prediction markets usually use one of two ways to organize trading. Real money markets (and some play money markets) use a dual auction. Other markets use a "market maker," which is described below. In a dual auction, traders can submit limit orders or market orders. Limit orders are offers to buy or sell at a specified price. These offers may or may not be accepted by other traders in the market. In contrast, market orders are executed immediately on the basis of previously posted limit orders.

The second major approach is to use some type of market maker system. The HSX market is one such example. In this type of market, buyers and sellers submit their orders to the market maker, which determines the prices using a predesigned set of rules. The advantage of this system over the double auction is a much lower barrier to entry for traders without experience (or training) in market trading.

There is at least one possible drawback to using a market maker system. Consider a trader who sees a favorable price in the HSX market and submits an offer to buy at the "current" price. The price observed by a trader is only a reflection of the last price at which trades took place. If there are a larger number of buy orders than sell orders placed within a short time period, the trader may end up paying much more than he or she expected once the market maker executes the set of concurrent buy and sell orders. This is a less important issue for securities like Moviestocks with a large range and a continuous outcome space. Such pricing uncertainty would be very problematic for prediction markets where the outcome range is narrow. For example, whether a contract is over- or undervalued may be a matter of a few cents in a prediction market forecasting popular vote shares of a very close election.

Prediction Markets of External Experts

In the wake of the intelligence failures surrounding the 9/11 attacks, the US

Department of Defense (DoD) undertook an extensive program of research on information systems to identify threats to national security. Managed through the Defense Advanced Research Projects Agency (DARPA), the FutureMap program was intended to develop market-based methods to provide accurate assessments of the likelihood of future events of interest to the DoD. Traders in these prediction markets would be independent experts with a variety of backgrounds. Ultimately, a description of a pilot study was incorrectly characterized in the media as a "terrorism market." Subsequent political opposition to the program led to its cancellation in 2003 (Surowiecki, 2004).

Unfortunately, there have been few subsequent prediction markets that have attempted to aggregate information from a varied set of experts. This is an important issue, since some prediction markets that rely on self-selected traders from the general public generate very accurate forecasts (i.e., US election markets) while the predictions in other domains are much less satisfactory (i.e., forecasts of movie box office results). This is an important area for future research.

One exception is a series of prediction markets using different types of medical personnel, for example, physicians, school nurses, and public health officials to generate short-term forecasts (two weeks ahead) of the severity of local outbreaks of influenza (Polgreen, Nelson, and Neumann, 2007). While this application is promising and may be useful in other areas of public health, it is still unknown whether expert traders can consistently predict low-probability, high-impact events (e.g., an unexpected early call for elections).

Preference Markets

Building on the success of prediction markets, a group of researchers at Massachusetts Institute of Technology (MIT) developed a market-based method for measuring consumer preferences for product concepts, attributes or even levels of a product attribute (*see* CONCEPT TESTING). Like a prediction market, participants in a preference market buy and sell securities. These represent, for example, alternative product concepts. Unlike prediction markets, however, there is no objective outcome that is used to determine payoffs to traders. Instead, the goal of the traders is to maximize the value of the portfolio of securities they hold at the end of a short trading session (usually about 15 minutes). The prices of the securities are determined by the traders' personal preferences for the various concepts and the traders' assessment of others' preferences for the concepts. Traders are rewarded based on the value of their portfolio relative to other traders. Incentives for MBA students participating in these markets have included gift certificates or small amounts of cash ($50).

To test the external validity of the results of the preferences, the researchers compared the market prices for a set of product attributes with two sets of preference measurements from traditional marketing research methods. The correlations were higher for the markets of MBA student traders (0.802 and 0.885) compared to those from markets where industry managers were traders (0.714 and 0.769). This suggests that preference markets can be useful in measuring preferences for a product concept, an attribute, or the level of an attribute (Dahan, Soukhoroukova, and Spann, 2009).

There are some potential limitations to using preference markets instead of traditional approaches such as conjoint analysis. The first is the lack of actionable segmentation of the market. A key advantage to conjoint analysis is the ability to segment the overall market with respect to differences in preferences for various attributes and levels of attributes. By design, the prices in a preference market can only provide an estimate of the consensus of the traders' preferences.

The second issue is that of trader selection. In public prediction markets such as the IEM election markets, traders are self-selected. Those who believe they possess superior information can choose to participate at the risk of their own funds. As noted above, the pools of traders need not be representative of the voting population for the market to produce an accurate forecast. In contrast, a preference market needs to confine its traders to a representative sample of consumers or a set of traders with a sufficient level of insight into the target consumers. In the case of using a representative sample of consumers, there is

no real advantage for the researcher in terms of sampling costs and effort between using a preference market or conducting a conjoint study. The results above show that even managers with industry experience do not necessarily have insights into the preferences of their consumers. This begs the question of how to recruit, screen, and motivate traders with a sufficient level of knowledge about the preferences of consumers who are not like themselves.

A final challenge for preference markets is to measure the relative importance of price versus other attributes. Traders in preference markets are not really making trade-off decisions as do respondents in a conjoint analysis study. Being able to put a dollar value on different levels of an attribute is one of the most important outputs of any conjoint analysis study.

A major advantage of preference markets over traditional methods of preference measurement is scalability and engagement of participants. Recently developed preference markets can quickly identify the clear winners and losers from a very high number of product concepts or attributes. Furthermore, traders are very positive about the experience.

New-product Idea Markets

Preference markets evaluate the relative attractiveness of different product concepts using consumers as traders. In new-product idea markets, the goal is to elicit and evaluate alternative new-product ideas using a company's employees as traders. As an idea-generating method, an Internet-based market has a number of important advantages. By using the Internet as the platform for interaction, an idea market can involve a large number of employees who are widely scattered across the globe. Since trading takes place over a number of days (or even weeks), participants can see ideas developed by others and create new options that improve upon earlier ideas. The prices generated by a market mechanism provide a continuous measure of the relative attractiveness of various proposed ideas.

An application of idea markets with one division of General Electric (GE) found that more ideas were generated and more employees participated than is typical for an idea generation exercise. There was a comparatively low correlation (0.43) between the market's ranking of an idea (based on the price of the associated security) and the average evaluation of the division's leadership team (LaComb, Barnett, and Pan, 2007). A subsequent idea market conducted for GE's Energy division resulted in immediate funding for the top two ideas and plans to patent several others (Spears *et al.*, 2009).

Key Challenges

Over the past few years, there has been increasing interest in using market-based methods to gather information. Researchers have developed various types of markets to tap into the collective wisdom of a firm's employees, its potential customers, or the public at large. Whether prediction markets use real or play money, the general public or expert traders, a double-auction or market marker, the results are consistent. The resulting forecasts have a high level of absolute accuracy and are often clearly superior to the next best alternative method. More recently, pioneering efforts have shown the ability of a market mechanism to elicit and evaluate new-product ideas from employees or measure consumer preferences with respect to product concepts, attributes, or attribute levels.

Despite these advances, there are still challenges to be addressed for managers interested in deploying these tools within their organizations.

Long-range forecast accuracy of prediction markets. The forecasts arising from the prices in prediction markets tend to have a very short-term focus. Market predictions of political candidate vote shares are determined on the eve of election day. The accuracy of some HSX forecasts is based on how well the prices on Friday afternoon (the day a movie opens in theaters) predict the first weekend's box office receipts (Pennock *et al.*, 2000). In neither of these applications would these forecasts be actionable. More research is needed on the ability of prediction markets to make longer range forecasts such as those available from more conventional (statistical) methods.

Motivating traders. One of the prerequisites in obtaining an accurate forecast from a prediction market is a trader pool that includes a number

of motivated traders with expert-level insights into the outcome of interest. Internal prediction markets can use real money incentives to motivate potential traders to share their valuable information with the organization through their participation in the market. However, the use of real money incentives raises some organizational barriers. Top managers may question the need to pay their managers to participate in a prediction market. The participating managers should already be supplying the firm with the best possible information. On the other hand, without incentives, it may be difficult to motivate a manager with valuable information to spend time learning how to trade in the market. In general, monetary rewards are generally modest to avoid any potential conflict of interest between a manager's trades in the prediction market and his or her ability to influence the outcome being forecast.

Expanding the use of preference markets. The preference markets reported in the literature focus on different types of durable products (cellular phones, crossover vehicles, laptop bags, bicycle pumps, computer games). However, the consumer traders in all of these markets are very highly educated MBA students. The trading market software required training and knowledge of double-auction markets (e.g., limit orders vs. market orders). Learning to operate in such a trading environment should be easy for MBA students but may be considerably more difficult for the general public. Whether preference markets can be successfully applied in other product categories or with less analytically inclined consumer populations is an important area of future development.

Summary

Leading companies including Google (Cowgill, Wolfers, and Zitzewitz, 2008), Best Buy, and GE are using various types of markets to access the collective "wisdom of the crowd" (Surowiecki, 2004). Whether the group is made up of employees or customers, there are new, market-based tools of discovery that leverage the reach and interactivity of the Internet to provide an engaging experience for participants and valuable information to organizations.

Bibliography

Chen, K.Y. and Plott, C.R. (2001) *Information aggregation mechanisms: concept, design and implementation for a sales forecasting problem*. California Institute of Technology working paper.

Cowgill, B., Wolfers, J. and Zitzewitz, E. (2008) Using prediction markets to track information flows: evidence from Google. www.bocowgill.com/GooglePredictionMarketPaper.pdf.

Dahan, E., Soukhoroukova, A. and Spann, M. (2009) New product development 2.0: preference markets. *Journal of Product Innovation Management*. (forthcoming).

Forsythe, R., Nelson, F., Neuman, G. and Wright, J. (1992) Anatomy of a political stock market. *American Economic Review*, **82**, 1142–1161.

Gruca, T.S., Berg, J. and Cipriano, M. (2005) Consensus and differences of opinion in electronic prediction markets. *Electronic Markets*, **15** (1), 13–22.

Gruca, T.S., Berg, J. and Cipriano, M. (2008) Incentives and accuracy issues in movie prediction markets. *Journal of Prediction Markets*, **2** (1), 29–43.

Hahn, R.W. and Tetlock, P.C. (2006) A new approach for regulating information markets. *Journal of Regulatory Economics*, **29**, 265–281.

LaComb, C.A., Barnett, J.A. and Pan, Q. (2007) The imagination market. *Information Systems Frontiers*, **9**, 245–256.

Pennock, D.M., Lawrence, S., Giles, C.L. and Nielsen, F.A. (2000) The Power of Play: Efficiency and Forecast Accuracy of Web Market Games. Technical Report, 2000-168. NEC Research Institute.

Polgreen, P.M., Nelson, F.D. and Neumann, G.R. (2007) Use of prediction markets to forecast infectious disease activity. *Clinical Infectious Diseases*, **44**, 272–279.

Rosenbloom, E.S. and Notz, W. (2005) Statistical tests of real-money versus play-money prediction markets. *Electronic Markets*, **16**, 63–69.

Spann, M. and Skiera, B. (2003) Internet-based virtual stock markets for business forecasting. *Management Science*, **49**, 1310–1326.

Spears, B., LaComb, C., Interrante, J. *et al.* (2009) "Examining trader behavior in idea markets: an implementation of GE's imagination markets. *Journal of Prediction Markets*, **3** (1), 17–39.

Sunder, S. (1995) Experimental asset markets: a survey, in *The Handbook of Experimental Economics* (J.H. Kagel and A.E. Roth), Princeton University Press, pp. 415–500.

Surowiecki, J. (2004) *The Wisdom of Crowds*, Doubleday, New York.

Jeffrey B. Schmidt

Levels of the Offering

Product planners should see their organization's offerings in the market at three levels, as shown in the Figure 1, as customers do: core customer benefit, actual product, and augmented product (Kotler and Armstrong, 2010).

Core Customer Benefit

At the most basic level, products and services are purchased to provide some perceived benefit(s) to the user. Marketers, designers, and product planners must determine "What is the customer *really* buying?" For example, a smart phone is more than a device that allows one to make phone calls, check email, access the Internet, and send text messages. Instead, it provides the freedom and the ability to connect to information and people from many locations. Similarly, a wristwatch might give the user the ability to be punctual for meetings and appointments, while certain watches might also communicate one's social status (e.g., Rolex) or the activities that one is interested in or participates in (e.g., running, aviation).

Actual Product

The actual product (or service) is how the core customer benefit is manifested or made available for consumption. Various features, design characteristics, and elements, the brand name, level of quality, and packaging, along with other attributes, comprise the actual product. In essence, the actual product is made from a collection of attributes (i.e., features, functions, and benefits) (Crawford and Anthony di Benedetto, 2008).

Augmented Product

The outer level in Figure 1, the augmented product, consists of additional products, services, and benefits that enhance/facilitate buying (such as credit and delivery terms), using, and disposing of the product (e.g., after-sales service, warranties, and technical

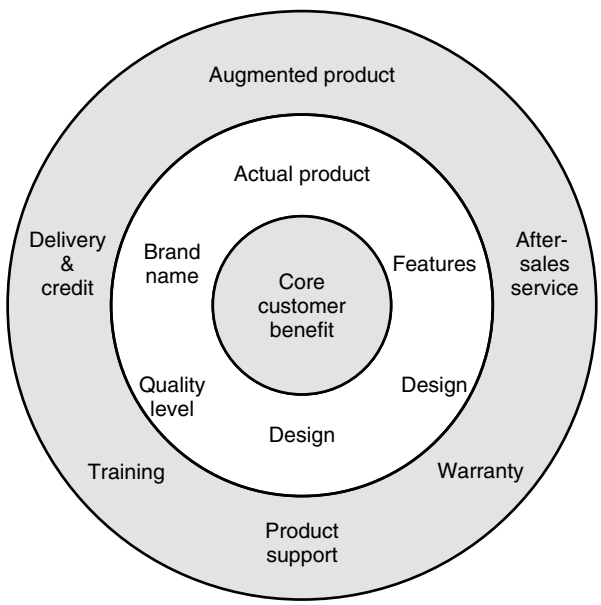


Figure 1 The whole-product concept. Source: Kotler and Armstrong, (2010).

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

assistance). Components of the augmented product offer ways in which to differentiate offerings when customers perceive little variance in the actual product or service.

Whole-Product Concept and Customer Value Proposition

VALUE PROPOSITION can be defined as what the customer gets, relative to competitors, compared to what he/she pays. Consequently, the whole-product concept and customer value proposition are related concepts. The whole-product concept equates to what the customer gets; that is, it is the entire bundle of attributes and benefits.

Bibliography

Kotler, P. and Armstrong, G. (2010) *Principles of Marketing*, 13th edn, Prentice-Hall, Upper Saddle River, p. 225.

Crawford, C.M. and Anthony Di Benedetto, C. (2008) *New Products Management*, 9th edn, McGraw-Hill Irwin, Boston, p. 132.

Kano model of customer satisfaction[1]

Barry L. Bayus

The Kano model of customer satisfaction is a widely cited approach for determining the types of customer needs (Kano *et al.*, 1984) (*see also* VOICE OF THE CUSTOMER; QUALITY FUNCTION DEPLOYMENT (QFD)). The Kano model adapts Fredrick Herzberg's ideas on the asymmetrical factors related to job satisfaction and dissatisfaction (i.e., job satisfaction is related to "motivators" like achievement, recognition, work itself, responsibility, whereas job dissatisfaction is related to "hygiene" factors like company policy, relationship with supervisor, work conditions, salary; Herzberg, Mausner, and Snyderman, 1959). Kano was working with the Konica camera company in the 1970s to develop some highly differentiated new products (Scholtes, 1997). Konica's sales and research groups discovered that customers only asked for minor improvements to the existing camera models. Kano, however, believed that really new innovations did not come from simply listening to what customers were verbally saying, but the development team had to develop a deep understanding of customer's real (latent) needs. Consequently, Konica staffers went to commercial photo processing labs to investigate the actual prints taken by customers. They found many mistakes and failures: blurry images, under and overexposure, and blank film rolls. Addressing these latent needs led to features such as auto focus, built-in-flash, and automatic film rewinding that are widely available in cameras today.

The key concepts in Kano's model are summarized in Figure 1. The horizontal axis indicates the degree to which a particular customer need is addressed in a (new or existing) product or service, ranging from completely absent to completely addressed. The vertical axis indicates how satisfied the customer is for a specific implementation of a customer need, ranging from delighted to disgusted. Within this two-dimensional space, three essential types of customer needs can be defined.[2]

The bottom curve, labeled *basic needs*, represents needs that are taken for granted and typically assumed by the customer to be met (i.e., these are needs that "must be" satisfied). "The car is safe, it uses regular gasoline, it can be parked in a standard parking space" are examples of basic needs for an automobile. These needs are the "order qualifiers" and thus must be completely satisfied; these needs are needed to simply be in the game. Completely meeting basic needs cannot greatly increase customer satisfaction, but if they are absent or below par customers will not react favorably.

The middle curve, labeled *performance needs*, represent needs for which customer satisfaction is roughly proportional to the performance exhibited by the product or service (i.e., these needs are "linear" in that "more is better"). For example, greater gas mileage and longer lasting tires for an automobile are typically preferred. These needs are frequently requested by customers during the course of traditional market research studies, and are usually associated with predictable product improvements.

The upper curve, labeled *exciting needs*, represent needs that the customer does not expect to be satisfied. Thus, if this need is completely addressed the customer is delighted but if not, the customer does not really care. These needs are the "order winners" for customers. For example, side airbags, global positioning systems, airless tires that never get flat for automobiles might be exciting needs today.

The underlying message of the Kano model is simple, yet powerful. Customer needs are dynamic in that an exciting need today will eventually migrate to being a performance need and will become a basic need tomorrow (e.g., automobile air conditioning was a delighter in the 1950s, but is a basic need today; more recently antilock braking systems and cup holders that were once exciting needs have become standard equipment in most automobiles). Thus, customer expectations increase over time and, consequently, firms must continually strive to better understand evolving customer needs in order to stay competitive. A video tutorial of the Kano Model is at http://www.kanomodel.com.

ENDNOTES

[1] This article is adapted from material in Bayus (2008).

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

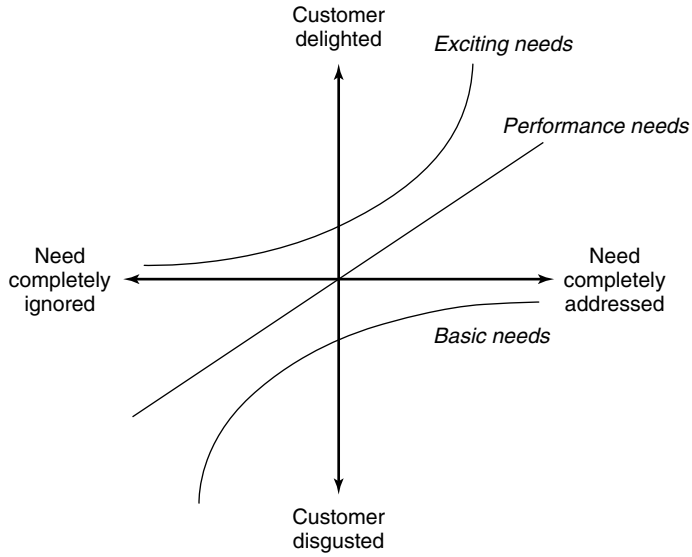


Figure 1 The Kano model.

[2] Other types of needs can also be defined based on the reverse of the curves in Figure 1, as well as needs for which the customer is indifferent (along the horizontal axis). See Center for Quality of Management (1993) and Matzler and Hinterhuber (1998) for a detailed discussion of methods to collect customer information that can be used to classify needs into these types.

Bibliography

Bayus, B.L. (2008), Understanding customer needs, in *Handbook of Technology and Innovation Management* (S. Shane), John Wiley & Sons, Ltd, West Sussex, pp. 115–141.

Center for Quality of Management (1993) Special issue: Kano's method for understanding customer-defined quality. *Center for Quality of Management Journal*, **2** (4), 3–36.

Herzberg, F., Mausner, B., and Snyderman, B. (1959) *The Motivation to Work*, John Wiley & Sons, Inc., New York.

Kano, N., Tsuji, S., Seraku, N., and Takahashi, F. (1984) Attractive quality and must-be quality. *Hinshitsu: The Journal of Japanese Society for Quality Control*, **14** (2), 39–48.

Matzler, K. and Hintehuber, H. (1998) How to make product development projects more successful by integrating Kano's model of customer satisfaction into quality function deployment. *Technovation*, **18** (1), 25–37.

Scholtes, P. (1997) *The Leader's Handbook*, McGraw-Hill, New York.

value co-creation

Gaurav Bhalla

Value co-creation refers to innovation practices jointly undertaken by a company and its consumers, both current and potential (*see* OPEN INNOVATION). The primary purpose of value co-creation is to create *unique and specific value for individual consumers*. Contrast this with traditional marketing approaches, in which value is created for *average consumers*, either in the total market, or in particular market segments.

A simple case study will help illustrate the dynamics of co-creation.

Imagine it is the early 1990s. A lady, let us call her Bella, experiences an increase in acidity and stomach pain each time she eats a major meal. She visits a gastrointestinal specialist (GI), who diagnoses her condition as GERD (gastroesophageal reflux disease), and prescribes her Zantac. Bella accepts the specialist's diagnosis, does not contest his expertise, asks a few clarifying questions and leaves.

Now fast forward to 2010. There is another lady, let's call her Chuki, who also experiences an increase in acidity and stomach pain, following each major meal. She too visits a GI, who also diagnoses her condition as GERD, but prescribes a stronger medicine, an acid suppressant called *Prilosec*.

However, Chuki is not willing to merely be handed a diagnosis and a prescription. She has many questions, wants answers, and proactively engages the specialist in a conversation about her condition (*see* VOICE OF THE CUSTOMER).

- She informs the specialist of friends who have similar symptoms, but were diagnosed with IBS (irritable bowel syndrome).
- She refers to an article discussing common stomach ailments that she has printed from WebMD in preparation for her visit.
- She has read on the Internet that proper amounts of stomach acid help digestion. She wonders whether taking a drug that suppresses stomach acid is a smart idea.
- What about alternative remedies – like natural cures and home remedies, shouldn't she try them first?

The brief case study illustrates the DNA of value co-creation. Chuki is accepting of the specialist's expertise, but not to the total subordination of her own opinions and knowledge. For her, value does not merely lie in receiving a diagnosis and a prescription. Her concept of value is broader. It encompasses information, knowledge, and the assurance that the specialist is treating her as a unique patient. What is more, she is willing to work with the specialist to this end.

The rise of the internet-connected, well informed, globally aware consumer, has significantly altered the balance of power between manufacturers and consumers. Today's consumers are unwilling to be passive recipients of mass-merchandised value. Instead, through a blend of assertion, activism, and participation in web conversations, today's consumers are demanding that companies shape their offerings to uniquely meet their specific needs, in specific situations, at specific points in time.

This shift in power, with consumers no longer merely being a source of demand at the end of a value chain, has fundamentally altered the way in which marketing is practiced and value is created (Prahalad and Ramaswamy, 2004). In today's world, value is co-created by a company and its many customers/noncustomers. The consumer is a fundamental resource and input in the value identification, creation, and delivery process, and the co-creation experience is an integral part of the total value offering.

The key building blocks of co-creation of value are (Bhalla and Lawrence, 2010)

- *Listening*. learning about consumers' experiences; their angst, frustrations, desires, and aspirations (*see* OPPORTUNITY IDENTIFICATION);
- *Sustaining value co-creation conversations*. meaningful conversations that yield the raw material for co-creation;
- *Experimenting and rapid prototyping*. to manage risk, improvise, and enable speedy value co-creation (*see* PROTOTYPE);
- *Execution*. only when co-created value is delivered can the next round of value co-creation be initiated.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

A few leading examples of co-creation of value follow. In the interest of space, web links are provided in lieu of detailed description.

- Nike + iPod allows gym fitness enthusiasts and runners to co-create their own fitness experiences and share it with similar others – www.nikeplus.com;
- You dream it, ASUS builds it, Intel inside – ASUS and Intel are combining forces to build the first community-designed laptop – http://www.wepc.com;
- Can innovation lead to evolution? Toyota believes it can. It is engaging millions of people to co-create value in domains like water, air, and energy – http://www.toyotawhynot.com.

Bibliography

Bhalla, G. (2010) Collective Voice: A Blueprint for Value Co-creation, Forthcoming.

Prahalad, C.K. and Ramaswamy, V. (2004) *The Future of Competition: Co-creating Unique Value with Customers*, Harvard Business School Press, Boston.

research & development

Jaideep C. Prabhu and Rajesh K. Chandy

What is R&D?

The Organisation for Economic Cooperation and Development (OECD, 2008) defines Research and Development (R&D) as "creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture and society, and the use of this stock of knowledge to devise new applications."

Research activities can be basic or applied. *Basic research* seeks to gain more comprehensive knowledge or understanding of a subject without specific applications in mind (National Science Board, 2010). Although basic research may not have specific applications as its goal, it can be directed in fields of present or potential interest. This is often the case with basic research performed by industry or mission-driven government agencies. *Applied research* seeks to gain knowledge or understanding to meet a specific, recognized need (*see* VOICE OF THE CUSTOMER). In business, applied research includes investigations to discover new scientific knowledge that has specific commercial objectives with respect to products, processes, or services (National Science Board, 2010). (*see also* SERVICE INNOVATION MANAGEMENT; PROCESS INNOVATION.)

Development is the systematic use of the knowledge or understanding gained from research or practical experience directed toward the production or significant improvement of useful products, services, processes, or methods (National Science Foundation, Division of Science Resources Statistics, 2006).

The main measure used to make comparisons of R&D activity across countries is gross domestic expenditure on R&D (GERD). This measure consists of the total expenditure on R&D by all companies, research institutes, and government and university laboratories within the country. (*see also* LOCALE OF INNOVATION.)

Who does R&D?

Two main sources of funding drive R&D within countries: public and corporate. Public funding drives R&D activity in government labs and research institutes as well as in public universities. Corporate funding drives R&D within firms. The extent of public versus corporate funding that drives R&D differs across countries. In some countries, most R&D activity is funded by corporations and takes place within in-house labs. For example, in Japan and South Korea, 77% of R&D is funded by corporations (National Science Board, 2010). In other countries, the government is still the largest spender on R&D and most of this activity takes place within state-sponsored networks of institutes and labs. For example, in India, over 80% of R&D is funded by the state (Evalueserve, 2008). The overall spending on R&D also varies significantly across countries. While most developed economies (e.g., USA, Japan, and Sweden) spend more than the OECD average of 2.2% of gross domestic product (GDP) on R&D, most emerging markets (Turkey, Brazil, India, and Mexico) spend less than 1% of GDP on R&D (OECD, 2008).

Where is R&D Done?

Historically, most R&D has been done in the developed economies of North America, Western Europe, and Japan (the so-called triad nations). Governments in these countries have spent large amounts of money on R&D to ensure leadership in defence. Likewise, the largest firms in these countries, which also count among the largest firms worldwide, have spent significant amounts on in-house R&D as a way to achieve long-term advantages over their global competitors. Indeed, China, India, and Taiwan are all now in the top 10 highest spending countries worldwide in overall (purchasing power parity adjusted) R&D expenditures (Duga and Studt, 2008). Moreover, many firms from triad nations have, in recent years, increasingly conducted R&D activities offshore, in emerging markets like China, India, Taiwan, Brazil, and Russia. For instance, a recent study finds that China has the fifth and India the seventh largest number of R&D centers of the Fortune 500 list of largest

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

firms worldwide (Tellis *et al*., 2009). This is particularly remarkable given that the trend toward offshoring R&D to emerging markets is barely two decades old.

R&D as a Concept and a Practice

R&D is a central concept in the vast academic literature on innovation (Thursby and Thursby, 2006). It is also a major instrument for policy makers striving to improve the competitiveness of their home economies and for managers seeking to improve the competitiveness of their firms worldwide (*see also* COMPETITIVE ADVANTAGE). While R&D is an important input into the innovation process, it does not guarantee that firms and nations that invest in it will naturally be more innovative. Indeed, a recent paper on innovation across nations finds that innovation and financial performance depend less on R&D spending per se, and more on firm-level cultural traits that enable firms to convert inputs into commercially valuable outputs (Tellis, Prabhu, and Chandy, 2009). Nevertheless, the tendency of firms, governments, and think tanks to monitor R&D spending is likely to continue in the future. The relative ease with which R&D spending can be measured (relative to innovation outputs) as well as its intuitive link with innovation outputs will ensure it retains its appeal and power for some time yet.

Bibliography

Duga, J. and Studt, T. (2008) Global R&D report. R&D Magazine, (September), G3–G18.

Evalueserve (2008) R&D Ecosystem in India, British High Commission, New Delhi.

National Science Board (2010) Research and development: national trends and international linkages, *Science and Engineering Indicators 2010*, Chapter 4, National Science Foundation, Arlington.

National Science Foundation, Division of Science Resources Statistics (2006) *Definitions of Research and Development: An Annotated Compilation of Official Sources* National Science Foundation, Arlington. http://www.nsf.gov/statistics/randdef/.

OECD Factbook (2008) Economic, Environmental and Social Statisstics, OECD, Paris.

Tellis, G.J., Eisingerich, A.B., Chandy, R.K., and Prabhu, J.C. (2009) Competing for the Future: Patterns in the Global Location of R&D Centers by the World's Largest Firms, unpublished working paper.

Tellis, G.J., Prabhu, J.C., and Chandy, R.K. (2009) Radical innovation in firms across nations: the pre-eminence of corporate culture. *Journal of Marketing*, **73** (1), 3–23.

Thursby, J. and Thursby, M. (2006) Where is the new science in corporate R&D? *Science*, **314** (8), 1547–1548.

value proposition

Shikhar Sarin

Introduction

Under increasing competitive pressure to keep costs down more and more, nowadays customers are focusing only on the price of the products they purchase. Without substantiation, a seller's claims of proposed benefits (i.e., value) are increasingly either ignored or dismissed outright as sales pitches and marketing charades (Anderson, Narus, and Rossum, 2006).

Definition

A *value proposition* is the sum total of all the benefits a seller will deliver to the target customer in a purchase (Best, 2004; *see also* VALUE CO-CREATION). An ideal value proposition is built around the key benefits sought by the target customer; where the target customer could be either individual organizations in business markets or segments in consumer markets (Best, 2004). A well-delivered value proposition can contribute significantly to the business strategy and performance of both the seller and the customer (Anderson, Narus, and Rossum, 2006).

Constructing a Value Proposition

A compelling value proposition is distinctive, measurable, and sustainable. Properly constructed value propositions force sellers to meticulously focus on the worth of their offering to their customer.

Building blocks of a value proposition. The foremost step in constructing a value proposition is identifying key value drivers for the customer (Best, 2004) or the value elements of a seller's offering (Anderson, Narus, and Rossum, 2006). One can identify value drivers either by asking customers directly what they value or by using approaches like conjoint analysis (*see* PRODUCT DESIGN; PRODUCT SPECIFICATIONS) to examine how customers make choices involving tradeoffs between combinations of different product attributes (Best, 2004). Similarly, from a sales perspective, sellers need to identify what they consider to be the key value drivers of their product offering (i.e., value elements). Once they have identified these value elements, the sellers need to ask themselves the following question: "From the customer's perspective, how do these value elements compare with those of the next best alternative?" Anderson, Narus, and Rossum (2006) suggest that on the basis of the answer to the above question, value elements of a seller's offering can be classified into three categories:

1. *Points of parity*. These are value elements that essentially offer the same performance or functionality as those of the next best alternative.
2. *Points of difference*. These are elements that make a seller's offering either superior or inferior to the next best alternative.
3. *Points of contention*. These are value elements about which the seller and the customer disagree regarding how their performance and functionality compares with those of the next best alternative (e.g., the seller considers a value element to be a point of difference, whereas the customer considers it to be a point of parity, or vice versa).

Types of value propositions. On the basis of these different value elements, sellers typically construct three kinds of value propositions (Anderson, Narus, and Rossum, 2006):

1. *All benefits*. These are value propositions in which the seller simply lists all the benefits they believe their offerings deliver to the target customer. This type of value proposition seeks to answer the customer's question, "Why should we purchase your offering?" An all-benefits value proposition requires the least amount of knowledge about the customer and the competitors, and the least amount of work in its construction. However, it often leads to benefit assertion by the sellers, where sellers create a laundry list of potential benefits. The problem arises from the fact that most of the benefits listed under this kind of a value proposition tend to be points of parity, which runs the risk of diluting the impact of genuine points

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

of difference with the next best alternative. Moreover, in all-benefits value propositions, sellers tend to claim advantages for features that provide little or no actual benefits for the target customers.

2. *Favorable points of differentiation*. These value propositions seek to differentiate the seller's offerings from the next best alternative. A favorable point of differentiation value proposition attempts to answer the customer's question, "Why should we purchase your product instead of the next best alternative?" Constructing such a value proposition requires detailed knowledge of competitive offerings. However, without knowledge of the customer, favorable points of differentiation proposition can lead to value presumption on the part of the seller. In other words, there is an assumption by the seller that the favorable points of differentiation must be valuable to the customers.
3. *Resonating focus*. These value propositions make the seller's offerings superior on the few key elements that matter the most to the target customer. A resonating focus value proposition seeks to answer the customer's question, "What is most worthwhile for us to keep in mind about your offering?" These propositions concentrate on one or two key points of difference that deliver the most value to the customer. Resonating focus value propositions are the hardest to construct, and require effort to gain insights into the value drivers of the customers.

To develop a compelling value proposition, the sellers needs to

- substantiate the value proposition (especially in financial terms);
- demonstrate customer value in advance; and
- document the value proposition in a manner that is consistent with the customer's priorities.

The reader is referred to Anderson, Narus, and Rossum (2006) and Best (2004) for further details on these aspects of a value proposition.

Bibliography

Anderson, J.C., Narus, J.A., and Van Rossum, W. (2006) Customer value proposition in business markets. *Harvard Business Review*, **84** (3), 91–99.

Best, R.J. (2004) *Market-Based Management: Strategies for Growing Customer Value and Profitability*, 3rd edn, Prentice Hall, Upper Saddle River.

first-mover advantage

Shikhar Sarin

Introduction

The concept of first-mover advantage (FMA), or pioneering advantage, has attracted considerable interest for over three decades. Given its theoretical and managerial significance, the topic has drawn attention from researchers from a variety of disciplines including marketing (Kerin, Varadarajan, and Peterson, 1992), strategy (Lieberman and Montgomery, 1988), technology management (Tushman and Anderson, 1986), and economics (Glazer, 1985). The present discussion first defines FMA, and then explores its advantages and disadvantages from three distinct perspectives: economic, behavioral, and empirical. Finally, the directions of future evolution of this research stream are presented.

Definition

Kerin, Varadarajan, and Peterson (1992) suggest that FMA generally refers to the widely held notion that being the first firm to pursue an opportunity is associated with long-term COMPETITIVE ADVANTAGE. A firm can achieve first-mover status in many ways, such as (i) being the first to produce a new product; (ii) being the first to use a new process; or (iii) being the first to enter a new market (Kerin, Varadarajan, and Peterson, 1992; *see also* OPPORTUNITY IDENTIFICATION). FMA is more complex than a simple order-of-entry effect; it incorporates a firm's ability to exploit and sustain any entry-related competitive advantage.

Literature Review

Academics and practitioners have long associated several advantages with being a first mover or a pioneer. Kerin, Varadarajan, and Peterson (1992, p. 42) note that pioneers preempt preferred "geographic space (locations), perceptual space (product characteristics space), distribution space (marketing intermediaries and shelf facings), and market segments (the largest and/or the most profitable)," which provide the first movers with superior resources relative to followers in a market (*see* PRODUCT POSITIONING). FMAs have been explored from a theoretical–analytical perspective, as well as an empirical perspective; with the theoretical–analytical approach further being categorized into economic and behavioral perspectives (Kerin, Varadarajan, and Peterson, 1992). These perspectives are briefly discussed next.

Economic perspective. Several advantages accrue to first movers from an economic perspective, foremost of which is a high barrier to entry for the followers in the market. These barriers to entry are based on a variety of factors. First movers have a faster learning curve from production and sales, and achieve greater economics of scale relative to late entrants. Additionally, pioneers are able to preempt scarce resources, make critical initial investments (i.e., developing distribution channels), and lock up critical relationships (i.e., suppliers, distributors) in binding contracts before the late entrants. This creates supplier-specific learning and contractual switching costs that are hard for followers to overcome (Kerin, Varadarajan, and Peterson, 1992; Szymanski, Troy, and Bharadwaj, 1995).

First movers are able to better protect their brands, products, and organizational infrastructure owing to ownership of proprietary technology, patents, copyrights, trademarks, and trade secrets (Lieberman and Montgomery, 1988). Control of critical technologies allows pioneers to exercise technological leadership over late entrants (*see also* LATER MOVER (NONPIONEER) ADVANTAGE). First movers not only benefit from greater name recognition and reputational effects (Kerin, Varadarajan and Peterson, 1992) but they also enjoy a higher switching cost from their products. Because of asymmetric information about product quality, buyers perceive an inherent risk in adopting lesser-known follower brands, leading to a high switching cost for the buyers (Lieberman and Montgomery, 1988; Szymanski, Troy, and Bharadwaj, 1995).

The economic perspective suggests that by creating significant barriers for the late entrants, the first movers are able to extend the time during which they operate with little or no competition. They use this period to not only skim the market for profit, but also maneuver themselves into

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

a dominant position in preparation for competition from late entrants (Kerin, Varadarajan, and Peterson, 1992). While the economic arguments in support of pioneering are impressive, the advantages are not automatic. Any of the potential economic advantages accorded by the order of entry are dependent upon the firm's ability to exploit and sustain such advantages (Kerin, Varadarajan, and Peterson, 1992).

Counterarguments from the economic perspective. Pioneering is risky. First movers not only incur the high cost of research and development, but they also bear the high potential for failure (Golder and Tellis, 1993). Late entrants can free ride on the investments in technology and market development made by the pioneer (Lieberman and Montgomery, 1988). Moreover, followers can learn from the mistakes of pioneers, and enter the market with superior marketing strategies (Golder and Tellis, 1993; *see also* LAUNCH STRATEGIES; GROWTH STRATEGIES). Often, inertia from incumbency constrains a pioneer's ability to respond to market changes effectively, and in a timely manner (Lieberman and Montgomery, 1988).

The economic perspective largely overlooks the fact that product–market contingencies are likely to moderate any sources of FMA (Kerin, Varadarajan, and Peterson, 1992). For example, demand uncertainty is likely to influence whether the first mover, or the followers, enjoy sustained competitive advantage in the market. Similarly, market uncertainty and technological turbulence increase the risk that followers may leapfrog the pioneer during shifts in technology or customer preferences.

Behavioral perspective. Compared to late entrants, first movers encounter lesser resistance from the early adopters and innovators of a product (*see* PRODUCT LIFE CYCLE; BASS MODEL). In cases where customers are unfamiliar with a product category, first movers play a critical role in shaping the preference structures of the consumer (Carpenter and Nakamoto, 1989). First movers not only get to influence how attributes are valued but also the combination of attributes that are considered "optimal." Thus, pioneers get to define a product category and the prototype in that product category in their favor; a standard against which all late entrants are measured (Carpenter and Nakamoto, 1989).

First movers enjoy a higher degree of consumer awareness than later entrants; they also find it easier to become a part of a customer's set of acceptable brands. This not only translates into higher product trials initially but also provides a basis for enduring sales advantage as the follower brands find it harder to penetrate the consideration sets of the customers (*see also* PRODUCT LIFE CYCLE; DIFFUSION OF INNOVATION; TAKEOFF; PRODUCT POSITIONING).

Counterarguments from the behavioral perspective. Kerin, Varadarajan, and Peterson (1992) note that there seems to be an inherent assumption in the behavioral perspective that the pioneer always offers a high-quality product, chooses the correct positioning, and pursues the right competitive strategy. This perspective further assumes that the late entrants are always reactive in their response and pursue a "me-too" strategy; since they lack the product offerings, business acumen, and organizational skills to attract the first mover's customers.

However, late entrants can learn from the pioneer's mistakes and develop superior marketing strategies (i.e., heavy advertising, more desirable positioning) that challenge the first mover's dominance (Golder and Tellis, 1993; *see also* LAUNCH STRATEGIES; GROWTH STRATEGIES; PRODUCT POSITIONING). Similarly, late entrants can choose to influence the consumer's preferences rather than simply reacting to them. Given that persuasive theoretical–analytical arguments can be found in support of, as well as against, FMA, many people look to the empirical evidence in this regard.

Empirical perspective. On the basis of a variety of data sources (i.e., PIMS database, sample surveys, and archival data), many empirical studies suggest a positive relationship between pioneering and market share, enduring sales advantage, firm survival, and profitability (see Kerin, Varadarajan, and Peterson (1992) for an excellent review; and Szymanski, Troy, and Bharadwaj (1995) for a meta-analysis), to the extent that the advantages of being "first to

market" have came to be regarded as conventional wisdom, or an empirical generalization.

However, other researchers have challenged this notion, reporting a negative relationship between order of entry and market share, long-term survival, and marketing and R&D expenses (Golder and Tellis, 1993; Kerin, Varadarajan, and Peterson, 1992).

This stream of research advocates firm growth and survival through following, rather than pioneering. A meta-analysis led Szymanski, Troy, and Bharadwaj (1995) to conclude that even though there is a positive association with market share, the overall empirical evidence in support of a direct effect of pioneering advantage is decidedly inconclusive.

Counterarguments from the empirical perspective. Several reasons have been suggested to help explain the mixed empirical support for pioneering advantage in the literature. Broadly, these inconsistencies can be attributed to the differences in the operationalization of pioneering, unit of analysis used (e.g., firm vs. strategic business unit (SBU) vs. brand/product level), sample heterogeneity and representativeness (e.g., survival bias; industry, business, or product type), and the measurement of performance (Lieberman and Montgomery, 1988; Kerin, Varadarajan, and Peterson, 1992; Szymanski, Troy, and Bharadwaj, 1995).

Past studies have operationalized pioneering in different ways. Some have operationalized it as being the first to develop a product or technology (not necessarily being the first to market), others have operationalized it as being the first to market (not necessarily being the first to develop), while still others have measured pioneering as being "one of the first" to develop a product or technology. This lack of consistency has led to contradictory findings because cases that might be considered as pioneering in one study would qualify as late entrants in a different study, and vice versa. In other studies, the distinction between first movers and followers may be confounded, leading to incoherent results.

Past studies have also been inconsistent about the level of analysis at which FMA has been examined. Some studies have examined the issue at the level of individual brand or product-market; others have examined it at a SBU or firm level. While it is acceptable to use brand and product levels synonymously as long as the brand consists of a single product, however, in most cases, firms and SBUs are associated with multiple products (*see* ORGANIZING FOR INNOVATION). While firms and SBUs may be first movers in one product-market, it is unlikely to be the case across multiple products. Conversely, the effects of pioneering are likely to be strongest at the level of the single brand or product-market level; they are likely to get increasingly diluted and harder to isolate at higher levels of analysis.

Concerns have also been raised about sampling issues related to some of the past studies on FMA. Many empirical studies suffer from survivor bias, where the data do not account for firms, SBUs, and products that failed. This raises questions about the representativeness of the samples, and the validity of the findings related to pioneering–performance relationships (Kerin, Varadarajan, and Peterson, 1992).

Kerin, Varadarajan, and Peterson also note that many empirical studies pool data across diverse businesses and industries. This creates a heterogeneity bias, because diverse product-markets are likely to have different levels of concentration, entry and exit barriers, and other structural characteristics that could systematically affect the performance differences between first movers and late entrants. Finally, performance variables along which FMA is evaluated (i.e., market share vs. sales growth vs. profitability vs. survival), and their method of measurement (i.e., when these measurements are taken) can lead one to very different conclusions (Kerin, Varadarajan, and Peterson, 1992; Szymanski, Troy, and Bharadwaj, 1995). Once again, past studies show little consistency in this regard, leading to conflicting findings.

Toward Contingency Models of First-Mover Advantage

The preceding discussion leads one to conclude that regardless of the perspective employed, the support for pioneering remains equivocal. This begs the question: Is there really an first-mover advantage? Emerging consensus suggests that

one of the leading causes for past inconsistencies may have been the simplistic main-effects approach employed by earlier studies. This approach focused on the direct link between order of entry and performance.

However, many omitted variables and contingent factors related to product, firm, and environmental characteristics can play an important moderating role in determining pioneering–performance relationship (Kerin, Varadarajan, and Peterson, 1992; Szymanski, Troy, and Bharadwaj, 1995). For example, it has been suggested that firm characteristics (e.g., size and incumbency), market strategy variables (e.g., breadth of the product line – *see* PRODUCT-LINE STRATEGIES; product quality, scale of market entry – *see* LAUNCH STRATEGIES; R&D reinvestment – *see* RESEARCH & DEVELOPMENT), and market environmental factors (e.g., market and technology growth rate – *see* GROWTH STRATEGIES; demand uncertainty) moderate the link between order of entry and performance.

Recent studies suggest that the pioneering–performance relationship may also be moderated by the characteristics of the innovation, such as radicalness of the innovation (i.e., continuous vs. discontinuous – *see* RADICAL INNOVATION), nature of the innovation (i.e., product vs. process innovation – *see* PROCESS INNOVATION), technological characteristics (i.e., network effects), and switching cost (Kerin, Varadarajan, and Peterson, 1992; Srinivasan, Lilien, and Rangaswamy, 2004; Varadarajan, Yadav, and Shankar, Forthcoming). Szymanski, Troy, and Bharadwaj (1995) find that a contingency approach to examining FMA has a significantly greater explanatory power, beyond the simplistic main-effects approach employed in the previous studies.

Conclusion

This discussion briefly explores FMA from economic, behavioral, and empirical perspectives. Each perspective offers persuasive arguments in favor of, and counterarguments against, pioneering advantage. The evidence in support of a simple main effect of FMA remains ambiguous. A more nuanced contingency approach to examining FMA in the future might resolve some of the contradictory and inconclusive findings from the past. Recent trends are encouraging, and the area offers a fruitful opportunity for further research.

Bibliography

Carpenter, G.S. and Nakamoto, K. (1989) Consumer preference formation and pioneering advantage. *Journal of Marketing Research*, **26**, 285–298.

Glazer, A. (1985) The advantages of being first. *The American Economic Review*, **75** (3), 473–480.

Golder, P.N. and Tellis, G.J. (1993) Pioneer advantage: marketing logic or marketing legend? *Journal of Marketing Research*, **30**, 158–170.

Kerin, R.A., Varadarajan, P.R., and Peterson, R.A. (1992) First-mover advantage: a synthesis, conceptual framework, and research propositions. *Journal of Marketing*, **58**, 33–52.

Lieberman, M.B. and Montgomery, D.B. (1988) First-mover advantages. *Strategic Management Journal*, **9** (Summer), 41–58.

Srinivasan, R., Lilien, G.L., and Rangaswamy, A. (2004) First in, first out? Network externalities and pioneer survival. *Journal of Marketing*, **68**, 41–57.

Szymanski, D.M., Troy, L.C., and Bharadwaj, S.G. (1995) Order of entry and business performance: an empirical synthesis and reexamination. *Journal of Marketing*, **59** (4), 17–33.

Tushman, M.L. and Anderson, P. (1986) Technological discontinuities and organizational environments. *Administrative Science Quarterly*, **31** (3), 439–466.

Varadarajan, P.R., Yadav, M., and Shankar, V. (Forthcoming) First-mover advantage in the internet-enabled market environment: conceptual framework and propositions. *Journal of the Academy of Marketing Science*. Fortcoming.

concept selection matrix

Jeffrey B. Schmidt

Introduction

Quality function development (QFD) is a technique that incorporates customers' needs and desires into various organizational processes such as engineering, design, marketing, and manufacturing (*see* QUALITY FUNCTION DEPLOYMENT (QFD)). The results of QFD are four matrices:

- Customer needs to performance measures (i.e., house of quality) (see Hauser and Clausing, 1988);
- Performance measures to features or solutions;
- Features or solutions to parts specifications;
- Parts specifications to manufacturing processes.

A variation on the second matrix, called the *Pugh concept selection*, was developed by Prof. Stewart Pugh of the United Kingdom. This variation attempts to evaluate different product concepts against the key performance measures of a "benchmark" or industry standard product with the goal of moving toward an "ideal" or near-perfect product concept. Pugh's concept selection matrix is a very effective for comparing concepts that are not refined enough for direct comparison using engineering requirements. This tool helps select the best design concept (or solution) from among alternatives, and, during the process, new concepts also may be generated.

Concept Selection Matrix

Concept selection entails both the processes of concept screening and concept scoring, each assisted by using a matrix (see Figure 1). For less complex solutions, only the concept screening matrix might be used, whereas more complex solutions might also require the use of the concept scoring matrix.

For both concept screening and scoring, the following six-step process is typically used by new-product design teams.

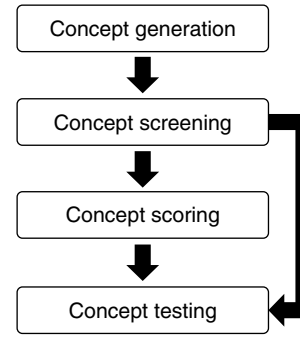


Figure 1 Concept selection.

- Prepare the selection matrix.
- Evaluate the concepts (design solutions).
- Rank the concepts.
- Combine and improve the concepts.
- Select the best concept.
- Reflect on the results and the process.

Each step of the concept screening is described briefly. Finally, concept scoring is discussed briefly, and a sample matrix is provided.

Step 1 Prepare the selection matrix: Using a matrix (on paper or in Excel, for example), the design team lists various possible concepts (or design solutions) on the horizontal axis and decision criteria on the vertical axis. These criteria are chosen on the basis of customer needs, and uncovered using VOICE OF THE CUSTOMER (Griffin and Hauser, 1993) and other marketing research techniques BRAINSTORMING. In addition, the needs of the organization (e.g., strategic importance, financial viability, fit with existing manufacturing techniques) should be considered.

Step 2 Evaluate the concepts (design solutions): The team chooses a benchmark (or reference) concept, which often is the industry standard or best-in-class product. The decision criteria for all concepts are rated against the benchmark concept.

Often a simple three-point scoring system is used (+, 0, and − for "better than", "equal to," and "worse than," respectively) by the team to compare each selection criteria for each concept. When available, objective measures (cost, engineering specifications) may be used.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Example of a concept scoring matrix.

		Opportunity 1		*Opportunity 2*		*Opportunity 3*		*Opportunity 4*	
	A	B_1	C_1	B_2	C_2	B_3	*C*	B_4	C_4
Decision criteria	Weight 1–5	Rating	Score	Rating	Score	Rating	Score	Rating	Score
Strategically aligned	4	9	36	5	20	7	28	1	4
Acceptable market size	3	6	18	8	24	2	6	10	30
Financially attractive	4	7	28	9	36	9	36	9	36
Competitively viable	5	8	40	6	30	9	45	9	45
Understood needs	3	8	24	5	15	3	9	9	27
Technically achievable	2	5	10	5	10	1	2	9	18
Total score			156		135		126		160

Source: Haines (2009).

Teams may arrive at a consensus rating or may use an average.

Step 3 Rank the concepts: After rating the concepts, the number of "better than," "equal to," and "worse than" ratings are summed for each concept. An overall score for each concept is computed by subtracting the number of "worse than" ratings from the number of "better than" ratings. Then, each concept is ranked relative to competing or alternative concepts.

Step 4 Combine and improve the concepts: Next the team decides if the concepts make sense and whether there is a way to combine two or more concepts into an even stronger one.

Step 5 Select the best concept: For the fifth step, the team picks one or more concepts for further refinement and exploration. In some cases, it might be advisable to pursue multiple concepts at this point until such time when one emerges as clearly superior.

Step 6 Reflect on the results and the process: The results are discussed so that the team members are committed to moving forward with the chosen concept(s). At this point, the team should reflect on the process so that subsequent concept scoring exercises might be improved.

Concept Scoring

The procedure for concept scoring is very similar to that for concept screening. The matrix in Table 1 shows the main differences. Rather than rating a concept as better or worse than the benchmark, the team gives each decision criterion a weight from 1 to 5 (e.g., from not very important to very important). Then, each opportunity is rated from 1 to 10 as in this example (e.g., from unacceptable to outstanding). Each organization can use the scales and anchor points that fit best or are most useful.

In Table 1, opportunity 4 is optimum (total score = 160), followed closely by opportunity 1 (total score = 156). These two might be continued forward. On the contrary, opportunities 2 and 3 (total scores = 135 and 126, respectively) might be discarded or at least held until the two more highly ranked opportunities can be developed further.

Bibliography

Griffin, A. and Hauser, J.H. (1993) The voice of the customer. *Marketing Science*, **12** (Winter), 1–27.

Haines, S. (2009) *The Product Manager's Desk Reference*, McGraw-Hill, New York, p. 285.

Hauser, J.R. and Clausing, D. (1988) The house of quality. *Harvard Business Review*, **66**, 63–73.

Product Development & Management Association Body of Knowledge, (2009) www.pdma.org.

Pugh, S. (1990) *Total Design*, Addison-Wesley, Reading.

Ulrich, K.T. and Eppinger, S.D. (2008) *Product Design and Development*, 4th edn, McGraw-Hill, Boston.

Warburton, D. (2004) *Getting Better Results in Design Concept Selection*. Medical Device & Diagnostics Industry, (2009) (1) http://www.devicelink.com/mddi/archive/04/01/006.html.

society, culture, and global consumer culture

Eric J. Arnould

Consumer Culture

Consumer culture can be defined as a "social arrangement in which the relations between the [lived cultural experience of everyday life] and social resources, between meaningful [valued] ways of life and the symbolic and material resources on which they depend, is mediated through markets." Consumer culture is a system in which consumption, a set of behaviors found in all times and places, is dominated by the consumption of commercial products. It is also a system in which the transmission of existing cultural values, norms, and customary ways of doing things from generation to generation "is largely understood to be carried out through the exercise of free personal choice in the private sphere of everyday life." Furthermore, consumer culture is also bound up with the idea of modernity, that is, a world "no longer governed by tradition but rather by flux," and in which "social actors who are deemed to be individually free and rational" hold sway (Slater, 1997, pp. 8–9). And finally, consumer culture derives from an economy in which value has been divorced from the material satisfaction of wants, and the symbolic value of goods takes precedence (*see* GLOBAL CONSUMERISM AND CONSUMPTION).

In consumer culture, predispositions toward social emulation, matching, and imitation expressed through marketplace choices are accompanied by a penchant for differentiation, individuality, and distinction also expressed through marketplace choices. Together, these motives drive a rapid turnover in goods and services. These dynamics are often thought to have been triggered by the purposeful social engineering of twentieth-century marketers, advertisers, and retailers and to have spread from roots in the fashion industry into all parts of social life (Featherstone, 1991, p. 115).

Four crucial aspects of consumer culture include

1. the pervasive and rapid circulation of commercial products, that is, things produced for exchange within a capitalist market, take priority over and above things redistributed by governments through the welfare state or exchanged among social groups through gift giving;
2. the relative independence of consumption activities from those related to production, and the growing power and authority this gives to some consumers over market dynamics;
3. changes in the relationships between different systems of production and valuation in society such that these are all increasingly interlinked and mediated by market values; that is, answers are sought to questions like "How much does it cost?", "How much will someone pay?", and "How much am I worth?"
4. the special importance given to the use of consumer goods in the allocation of individual status, prestige, perceived well-being, and quality of life (Lury, 1996, p. 4).

Consumer culture is influenced by agents who work directly in the market economy as managers, marketers, and advertising "creatives"; by independent "brokers" who analyze and criticize consumer products; by cultural intermediaries such as media figures (e.g., movie and television stars, celebrity chefs, religious broadcasters, public intellectuals, politicians, etc.) who sculpt and disseminate attractive models for consumption behavior; and by dissidents who initiate alternative responses to the mass-consumption system, responses that are typically reappropriated into the global market as differentiated, niche products.

Global Consumers

The consolidation of scientific economics and modern market institutions took place in the eighteenth century, which was also when the social role of the consumer, combining traits of hedonism and rationality, was distinguished from the putatively wasteful and irrational elite of the ancien regime. The early twentieth century consumer was a mass-market consumer, whereas today's consumer is characterized by a general emphasis on individual style, paralleling the customization and niche marketing that has

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

overtaken the economy. The tendency within consumer culture today is that consumers' lifestyles no longer require coherence; marketers and cultural intermediaries (fashion, entertainment) cater to and expand the range of styles and lifestyles available to global audiences with little emphasis on cultural authenticity or prior tradition (Featherstone, 1991, p. 26; *see* GLOBAL CONSUMERISM AND CONSUMPTION).

And just as the consumer was theorized into existence by the economic philosophers of the eighteenth century, and turned into the linchpin of twentieth-century economies by economic policy makers and Madison Avenue advertisers, so the consumer continues to be recreated on the global stage. For example, Dávila (2001) carefully traces the evolution of Spanish language media in the United States, as well as advertising specifically designed for Latinos. The dynamics of these initiatives prior to the 1980s often involved promotions aimed at Latin American countries that were then adapted and transplanted to the United States. Cuban advertising entrepreneurs and cultural capital and networking links generated by the Cuban-American ethnic economy proved central to the development of Latino media. She demonstrates that an emphasis on stereotypical "traditional" family values is at the crux of Hispanic marketing strategies, and the constitution of the Hispanic consumer, by these media. She notes how important uniting Latinos across difference of nationality, class, color, and political ideology has been in building an image of the Latino market for corporate clients (*see* PRODUCT ETHNICITY; GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES).

Further, Cayla and Eckhardt (2008) investigate how Asian brand managers forge new webs of interconnectedness through the construction of a transnational, imagined Asian world. Some brand managers are creating regional brands that emphasize a putatively common experience of globalization, evoke a generic, hyperurban, and multicultural Asian experience, and contribute to the creation of an imagined Asian as urban, modern, and multicultural (*see* GLOBAL BRANDING: THREE KEYS FOR GLOBAL BRAND SUCCESS; GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES).

Finally, Mazzarella (2003) has documented in greater detail how Indian marketing professionals marshaled ideas about Indian cultural distinctiveness along with stereotypical local notions of tradition and modernity to create a differentiated image of the Indian consumer to market to multinational corporations (*see* BASE OF THE PYRAMID MARKETS: CULTURE INSIGHTS AND MARKETING IMPLICATIONS; RESEARCH & DEVELOPMENT).

Global Consumer Culture

Global consumer culture is driven by the extension of market principles into every corner of the planet. Global consumer culture is a desired ideal for many new consumers, and at the same time, a contested foreign ideology for some political and religious groups, and a social process that is continuously evolving through shifting business relationships, competition, technological flows, and media expressions.

More specifically, the spread of consumer culture globally is conditioned by the global flow of distinctive cultural resources including what Appadurai (1990) refers to as *financescapes*, *technoscapes*, *ideoscapes*, and *mediascapes*. Financescapes refers to global flows not only of capital and currency but of various financial instruments that facilitate markets (*see* INTERNATIONAL PRICING OBJECTIVES AND STRATEGIES). The derivatives market, the remittances business, and the microfinance revolution are all examples. Remittance businesses have facilitated both the global movement of workers and also the transfer of consumption resources to the global south. The microfinance revolution has also driven market models of finance and consumption into the hands of the global poor.

Technoscapes refers to global flows of informational and mechanical technologies of both simple and complex varieties across borders. Integrated global supply-chain management at one scale of complexity (*see* DESIGNING A GLOBAL SUPPLY CHAIN: OPPORTUNITIES AND CHALLENGES), and the global cell-phone revolution at another are cases in point. Global

consumer culture also depends upon global media to create a sense of global identification and memory without which any cultural identity is incomplete. Thus, televised global consumption spectacles such as the millennial celebrations, and the quadrennial World Cup and Olympic festivals are implicated in the spread of consumer culture (*see* DIGITAL MEDIUM AND GLOBAL MARKETING).

In the global mediascape, privatized commercial media empires (e.g., NewsCorp) now coexist with decentralized and fragmented small media, diverse in economic and social organization such as the cassette music and video cultures that have become ubiquitous in the global south. Media of many forms play crucial roles both in shaping national imaginaries including national consumer prototypes such as the prototypical Indian, "Asian," or Chinese consumer (Dávila, 2001; Cayla and Eckhart, 2008; Mazzarella, 2003) mentioned above (*see* INTERNATIONAL ADVERTISING – IS THERE STILL A STANDARDIZATION VERSUS LOCAL ADAPTATION DEBATE?).

Ideoscapes refer to the ideologies of states and other movements that contend for popular legitimacy and authority. Globalization and standardization of modern consumer culture around the world have, in turn, stimulated localization and heterogeneity of demand as well as contestation and resistance to global companies and brands (Featherstone, 1991). Ideoscapes then include such things as models of consumer culture carried by global media; various reactions such as the anticonsumer No Logo movement, the rejectionist stream of Islamic fundamentalism that views film, music, fashion, and other trappings of consumer culture as anathema; and, more reformist movements like the local food and Slow Food movements (*see* FORCES AFFECTING GLOBAL INTEGRATION AND GLOBAL MARKETING).

Globalization also means that cultural encounters proliferate through these earthscaping processes, which has led to an increasing interest in identity-constructive processes through consumption. Thus, global consumer culture involves the globalization of desires; of the responsibility to seek an individual sense of self through material symbols; the need to emulate; the attraction of a market-mediated material world; of homogenized images of the good life (Clammer, 1997, p. 14); and an experience of fragmentation of social life on the receiving end of globalization that fuels idealist and rejectionist reactions (*see* GLOBAL CONSUMERISM AND CONSUMPTION).

Structures of common difference. Globalization is accompanied by a heightened sense of disjuncture and disorder, as rapid earthscaping has upset and displaced relationships between the economic center and periphery of the global economy, and between local cultural forms and cultural contents. These earthscaping processes and the novel technologies, media, ideologies, goods, and relations they carry "nevertheless often remain closely tied to professional, political, and economic interests that have important stakes in mobilizing and regulating global markets." These interests, "even as they capitalize on the proliferation of 'cultural difference,' also demand that such cultural difference be rendered manageable as content within globally reproducible (and thus marketable) forms and genres" (Mazzarella, 2004, p. 351; *see* FORCES AFFECTING GLOBAL INTEGRATION AND GLOBAL MARKETING; MARKET ENTRY AND EXPANSION). This has led to the identification of what have been called *global structures of common difference* in global consumer culture.

The creolization of consumption patterns that combine elements of local and foreign consumption traditions through globalized institutional forms is ubiquitous within global consumer culture. This is the diffusion of "structures of common difference" (Wilk, 1995). Thus, global fast-food chains reproduce an institutional formula across many cultural locations. McDonald's is present in most local markets, but often undergoes significant alteration in what is offered and how. Thanks to local modifications of the McDonald's formula, the Jolibee fast-food chain at one time enjoyed a 59% market share in the Philippine fast-food market. Similarly, traditional Turkish fast food called *kepab*, has experienced a revival as Western-style fast-food outlets have penetrated the Turkish market. And Turkish entrepreneurs in Germany offer Turkish *shish kepab* in McDonald's-like settings.

These choices allow Philippine, Turkish, and Turko-German consumers to express varying degrees of cultural conformity and cultural difference through their fast-food choices (*see* INTERNATIONAL FRANCHISING).

Similarly, the beauty pageant has become a popular consumer global celebration. In Belize in Central America, some pageants judge contestants according to Euro–American ideals of beauty; others celebrate Caribbean values of respectability and reputation and express tensions between different ethnic groups that hold different value orientations (Wilk, 1995). Soap operas, comic books, and musical forms provide other examples of global structures of common difference. The Hong Kong, Hollywood, and Bollywood industries provide three intertwined global institutionalized forms (Stearns, 2006, p. 150). And scholars have pointed out that Christmas has become the first global consumer holiday, a structure of common difference that accommodates dramatic differences in cultural content within a shared seasonal and aestheticized institutional format.

Brands. Brands have become a ubiquitous structure of common difference in global consumer culture: the Coca Cola logo and Nike Swoosh are brand symbols that trigger myriad responses; their cognitive salience and ability to arouse passion are undeniable (*see* GLOBAL BRANDING: THREE KEYS FOR GLOBAL BRAND SUCCESS). When people demonstrate against the inequities of globalization, they use brands such as Coke or McDonald's as symbols of corporate power. Brands have become cultural forms; they encapsulate ideas about the way people should live, look, and think. Branding is a specific form of communication that tells stories in the context of products and services, addresses people as consumers, and promises to fulfill unmet desires and needs. In other words, branding is a specific symbolic form, a particular way of talking about and seeing the world.

"The rise of a global culture does not mean that consumers share the same tastes or values. Rather, people in different nations, often with conflicting viewpoints, participate in a shared conversation, drawing upon shared symbols. One of the key symbols in that conversation is the global brand" (Holt, Quelch, and Taylor, 2004, p. 70). Global brands are most often associated with a quality signal that is important worldwide; global brands set a standard. Hence, global companies are advised to compete aggressively on quality signals while addressing consumers' skepticism about them. But global brands simultaneously create openings for local brands that convey enduring local meanings. Global brands also convey a myth of global cosmopolitanism to which many consumers worldwide aspire. Therefore, global companies are advised to associate global myths of individual independence, modernity, and self-actualization with their brands. Finally, global brands and global companies "wield extraordinary influence, both positive and negative, on society's well-being. Many consumers expect firms to address social problems linked to what they sell and how they produce and distribute their products and services" (Holt, Quelch, and Taylor, 2004, pp. 71–72). Thus, to improve their global image, firms are advised to undertake initiatives that benefit stakeholder communities (*see* MARKETING'S CORPORATE RESPONSIBILITY).

Regional Manifestations of Global Consumer Culture

OECD countries. The Organisation for Economic Cooperation and Development (OECD)-countries represent the most-developed market economies in the world, primarily, the Euro–American nations. In these countries, there is a steady multiplication of purchase opportunities and the extension of consumer orientations into areas of activity that used to lie outside the consumer domain, such as politics, higher education, art, place, and even interpersonal relationships. The center of the culture of consumption is marked by an oversupply of goods of all sorts including symbolic goods leading to a tendency toward cultural disorder and destructuring (Featherstone, 1991, p. 13). Individuals are encouraged to adopt a nonutilitarian attitude toward commercial goods and services and to carefully choose, arrange, adapt, and display goods to make stylistic statements that express the individual, presumably, authentic identity of the owner (Featherstone,

1991, p. 114), leading also to new consumer engagements in value-creation processes.

Four other trends can be noted. First, OECD countries are moving toward a dematerialized service economy (*see* SERVICES GLOBALIZATION). Second, trends toward recreational and luxury shopping have grown. Third, while elite consumers in other cultures have swelled the ranks of world tourism in recent years, North America and Europe have traditionally constituted large tourist markets. The motives that inspire tourist consumption are diverse, ranging from nostalgic visits to historic towns and villages, to escapist entertainment at Mardi Gras, to romance at a Club Med or on a cruise, to extraordinary adventures through African safaris, Himalayan trekking, or Outward Bound trips.

Fourth, is the advent of sustainability in consumer behavior that takes into account environmental concerns, that is, waste and resource depletion; social concerns, that is, "ethical" consumption, buycotts and boycotts, progressive human resources policies; and the efficiency and efficacy of firm operations themselves (*see* GLOBAL MARKETING ETHICS; EMBARGOES AND SANCTIONS). Negative (resistance) and positive (ethical buying) forms of political consumerism now seem to concern a wide sector of the population. Recent survey data on Europe, for example, show that over 30% of the population of Denmark, Great Britain, Germany, Italy, Norway, and Portugal are boycotting products for political reasons, choosing specific items because of their ethical or environmental qualities, or participating in the activities of consumer-oriented associations (Sassatelli, 2006, p. 221). Ethical consumerism is being reframed as consumer freedom and as a duty.

Consumer culture in Japan, the BRIC (Brazil, Russia, India, and China) countries, in the countries that were formerly part of the Soviet bloc, and in Africa (see below) has been "heavily foreign, a clear import, even as it appealed to both new and traditional interests. Foreignness, in turn generates three reactions. The first involves the appeal of the strange and ... the modern ... second, foreignness prompts resistance, in the name of customary but also newer national identities ... Third ... consumerism is appropriated, becoming as 'natural' as it was in the West" (Stearns, 2006, p. 81).

Japan and the Asian "Tigers". As one of the wealthiest nations in the world, Japan, representing a huge market, quickly adopted a culture of consumption after World War II. The direction of Japan's economic progress since World War II foreshadowed similar changes in other countries of the Pacific Rim, the so-called "tigers." While, Japan shares many historical values with other East Asian countries, export of finance-, media-, techno-, and ideoscapes originating in Japan have been instrumental in fueling the expansion of consumer culture in this region (*see* STRATEGIC EXPORT MARKETING–ACHIEVING SUCCESS IN A HARSH ENVIRONMENT).

A characteristic of Japanese consumers is that they are highly informed and aesthetically sophisticated. They are actively engaged in creating their sense of identity in terms of gender, age, and lifestyle (Clammer, 1997, p. 12). Further, there is less Puritanism (i.e., Protestant asceticism) and no moral condemnation attached to acquiring material goods in Japanese culture (Clammer, 1997, p. 14). Two interesting features of consumer culture in Japan are first, the association of high quality and price, hence a desire for brand-name goods, and the following of expensive fads; and second, the very high rate of discard and replacement. This reflects the fact that variety or originality are devalued in Japanese consumer culture relative to newness or up-to-dateness (Clammer, 1997, p. 24).

Japanese household budgets show interesting characteristics. Budget proportions spent on necessities, furniture, and household goods have declined, while amounts spent on leisure and luxuries have increased although there is still a high level of savings reflecting Japan's particular adoption of consumer culture historically, and such things as the high costs of housing and education. Gender differences are also significant, with rapid increases in women's discretionary expenditures. *Depāto* – department stores – which have gone a long way in commercializing Japanese cultural values and domesticating foreign consumer tastes, are also among the biggest promoters of "traditional"

Japanese gift giving – a large, distinctive, and socially important category of consumer expenditure in Japan (Clammer, 1997, p. 18). Japanese gift giving is a commercialized form of modern intimacy, a form that creates bonds without much moral substance, by contrast with China where commensalism is an important basis of social bonding (Clammer, 1997, p. 19).

BRIC countries. BRIC literally refers to Brazil, Russia, India, and China, but may include a number of countries with rapid economic development and growth in consumer culture, thus including countries like Turkey (*see* EMERGING MARKETS). In all these countries, the emergence of powerful local brands such as the Murat, Mahindra, and Tata automobile brands in Turkey and India, respectively, and large local consumer markets is a big part of the story of global consumer culture in the twenty-first century. Another part of the story is the persistence of large numbers of subsistence consumers at the base of the economic pyramid, who live on at best a few dollars a day and whose needs and aspirations have only recently begun to register on global marketers.

In China, the spread of consumer culture has been fostered by the existence of strong consumer interest before the nineteenth century as part of an urbane, secular worldview. The appeal of "an almost fantasy-like modernity" that promises some release from "customary hierarchies and constraints" (Stearns, 2006, p. 109) has also fostered the recent expansion of consumer culture. Finally, economic reforms post-1979 have diminished state power dramatically and freed many consumers from political strictures on consumption.

The World Bank estimates that 250 million to 300 million people have climbed out of poverty since China adopted economic reforms. One big story is the growth of a consumer middle class in China. In 2005, at least 4.5 million had a disposable income in excess of US$30 000 (Latham, 2006, p. 9). Characteristic of the effects of consumer culture everywhere, consumer culture is fueling the emergence of new "disjunctures and differences" in Chinese consumption practices (Latham, 2006, p. 3), new scope for Chinese citizens, especially women, to express their personal taste, ideals, and values as against older collective forms.

As in other places, distinctive characteristics of consumer culture are emergent in China. Among these are the idea of consumption as a palliative to continued tight state control of political freedom and the media; the articulation of various local understandings of Chinese history and character in its branding practices, and the role of consumer goods in vitalizing China's gift economy and, in particular, the web of interpersonal relationships often referred to as guanxi (Latham, 2006; *see* INTERNATIONAL RELATIONSHIP MARKETING). Others have commented on China's one-child policy in creating the more hedonically oriented Chinese consumer of today. The speed of economic development and the success of the one-child family policy has fomented a consumer revolution across all social strata. The claims of Chinese children on their family's financial resources are larger and more uniform across economic strata than for children elsewhere.

Much continuity persists in Chinese consumer culture as well; one example has to do with the place of food in society. As Mintz (2007) evocatively writes,

> It is at table that children learn to become adults; at table that babies meet their grandparents; at table that people display their civilization and communicate it. To watch the giver of a restaurant banquet—some paterfamilias welcoming the family of a son's fiancée, celebrating a grandchild's birth, or just treating friends—is to get a sober lesson in etiquette, self discipline, and joy.

In India, the growth of a middle class representing 30% of the population, or 250 million people is the big story of the early twenty-first century. Still, it is likely that no more than 100 million or so of this middle class represents a truly consumerist segment. As in China, participation in consumer culture is associated with cosmopolitan values like modernity, democracy, and even liberation. Nonetheless, the majority of the larger group remains committed to a savings orientation; Indians tend to save 25% of their incomes and place their savings in cash accounts rather than in financial instruments.

Across the population, about half of all discretionary expenditures are for important family ceremonial activities (Srinivas, 2008) rather than for the satisfaction of individual desires, indicative of the enduring role of family in structuring consumption practices. On the other hand, there is a huge aspiring class, perhaps as many as 260 million, striving to attain middle-class status, with maybe 15 million joining the middle classes each year (Srinivas, 2008, p. 11) who may engage in bouts of conspicuous spending. The Indian middle class is also divided along the fault lines of geography, especially the north–south divide, education, profession, especially differences between the private and public sector, class, and caste.

Researchers have found that the vast masses at the base of the Indian economic pyramid are also affected by the spread of consumer culture. "Increasing desires to consume branded goods that are advertised through television is . . . a consistent and recurring theme." Moreover, "intertwined cultural processes of conspicuous consumption, normative change [imposing a link between consumer goods and morality], and [interpersonal] competition" mark narratives of low-caste Indian consumers. They reflect an increasingly consumerist content of Indian media that depicts the mythic lifestyles of the rich and famous (Varman and Belk, 2008, p. 236, 237).

Russia and the former Soviet bloc. In Russia and other former Soviet bloc countries, the introduction of consumer culture has been inflected by the "debate about whether West should be model or pariah" (Stearns, 2006, p. 86). "Hesitations" rooted in the Soviet experience and Eastern Orthodox religious traditions persist (Stearns, 2006, p. 91). In all these states, there is a legacy of an effort to build an alternative to Western consumerism that emphasized workers' vacations, and workers' rights rather than consumer service. Further, there was a tradition of elite consumption that was both unacknowledged and secretive, a tradition that has reemerged in the spectacular nouveaux riche consumption behavior of the "new Russians," who profited from the Soviet Union's transition to market capitalism.

Latin America. In Latin America, a widespread class of wealthy landowners and a few merchants and mine owners provided a spur to consumerism in the nineteenth century and a model of European, and secondarily North American imitation that continues to color consumption, particularly, in countries like Argentina. Another characteristic of Latin America is the influence of American chain stores, first Sears Roebuck in the 1920s and later Wal-mart and some European chains in the 1980s and 1990s that began to extend Euro–American models of consumption to middle-class buyers. The middle classes that have aspired to Euro–American consumption standards have grown dramatically, though erratically, in Brazil and other Latin American countries since the Second World War. Of course, Latin America also encompasses dramatic cultural diversity that colors consumer preferences; in indigenous communities, sometimes, new consumer opportunities fuel cultural ideals that are a legacy of reciprocal social relations and new status hierarchies simultaneously. Latin America's centuries-long legacy of economic reliance on raw materials exports has produced a huge underclass of subsistence consumers, many of indigenous backgrounds, who struggle to enjoy basic consumer goods. These consumers are increasingly targeted by multinationals with custom-tailored products, creating new sustainability challenges (*see* EMERGING MARKETS). Finally, Latin America has also become an important source of exported consumer culture through such things as food (especially Mexican and, more recently, Argentinean culinary styles), music, and dance styles (Mambo, Tango, Meringue, Samba, etc.), and traditions of sumptuary seasonal consumption associated with Carnival and other traditions (Stearns, 2006, pp. 112–114).

Africa. Four intertwined themes concern Africa's engagement with consumer culture. First, in the nineteenth and twentieth centuries, as Africa was incorporated into the global economy through the sales of raw materials or labor to western interests, consumption became a practice allowing some to emphasize individual identity at the expense of forms of identity based on kin groups such as lineages and clans.

Imported consumer goods from Europe and, more recently, from Asia provide vehicles for constructing "modern" cosmopolitan consumer lifestyles. Today, exotic consumer goods from the metropolitan countries are used more widely to convey status and prestige. Cars (see the "mama Benz" of Togo or the wa-Benzi of East Africa), clothing, and buildings, both private homes and public buildings such as mosques, have become important indicators of status (*see* CONSUMER WORLD-MINDEDNESS).

Second, longstanding precolonial consumption practices typically remain strongly rooted, even as contemporary practices associated with the colonial and postcolonial experiences are layered over them. Thus, the share of household expenditures on food remains relatively high even at higher levels of income. The likely explanations for this pattern are relative poverty, the weakness of formal financial investment opportunities, and the normative importance of social investments in networks of kin, credit, and cult. Consumption of traditional products has, in many cases, persisted and, indeed, thrived. Fish from the inner delta of the Niger and the east African lakes, livestock and hides produced in semiarid regions, salt and dates produced in desert oases; kola nuts form the Guinea coast, and khat in the Horn region are some examples. Kola nuts, widely traded in West Africa, continue to play an important role in symbolizing hospitality, marital alliance, political affiliation, and ritual obeisance. Among the poorer consumer segments, motivations driving consumption choices are achieving increased family solidarity or social standing within local prestige networks, not individual consumer utopia.

Third, creolized consumption is ubiquitous. For example, there has been an explosion of fast food in African countries. Traditional fast food is mainly prepared with indigenous ingredients and to conform to local taste. "Foreign" models are more diversified, as urban populations embrace consumption patterns broadly similar to cosmopolitan ones under the impetus of resource availability and following "modern" trends. Proliferating fast food can be found in most of the larger African urban centers. Creolized clothing consumption may reach its apogee in the behavior of Congolese *sapeurs* who participate in elaborate rituals of conspicuous consumption of French fashions that incorporate traditional ideas about the accumulation of life force. Africans also have become avid producers and consumers of creolized music and video products, incorporating local, South Asian, and New World sources.

A fourth trend is the African consumers' imaginative engagement with alternative modern worlds that are different from their own but also without the ideological and political baggage associated with the west and the former colonial powers. The Islamic Middle East and Hindu South Asia offer alternative models of consumption.

The Islamic Middle East. The interaction between Islam and consumer culture represents interplay between ongoing spiritual and nationalist values and the new lures of displays of materialism. This occurs against a backdrop of public debate over secularism represented by countries like Turkey, Lebanon, Morocco, and Egypt, and Islamic religious nationalism represented by countries like Saudi Arabia, Iran, and the Sudan. A strong rural–urban, rich–poor divide, and between those who have benefited from oil revenues and those who have not, colors attitudes toward consumption (Stearns, 2006).

Huge numbers of internal migrants from rural villages to urban slums also find their consumption choices suspended between the expression of secular and religious values. Thus, poor consumers illustrate three modes of acculturation: in some cases, migrants reconstitute their village culture in the city, shutting out the consumerist ideoscape; or they collectively pursue the consumerist ideoscape as a myth through ritualized consumption; or they give up on both pursuits, resulting in anomic results for identity. Meanwhile, elite consumers in the Islamic Middle East are avid consumers of Western luxuries and middle-class consumers are often avid consumers of Western fashion and media. Among middle-class Islamists, Islamic fashions in dress, and gender-segregated options in shopping and luxury vacations have grown dramatically in the past 20 years (Üstüner and Holt, 2007; *see* PRODUCT ETHNICITY).

Resistance

Critique of and resistance to the spread of consumer culture is as old as the emergence of consumer culture in the eighteenth century. Systematic critique of the institutional bases of consumer culture has been offered by social theorists that highlights the alienating, dehumanizing effects of materialism; others have commented on the envy, possessiveness, and nongenerosity that commitment to consumer culture sometimes entails (*see* GLOBAL MARKETING ETHICS). Some classic expressions of resistance to consumerism have declined in recent years in the OECD countries although the critique of "irrational" lower class consumption or women's alleged frivolity persists in newer consumer cultures. Global consumer culture has, nonetheless, fostered a virulent backlash. Religious, environmentalist, nationalist, and anticorporate critiques have emerged with considerable vigor. Many of these forms of resistance appear to be motivated by a global sense of anxiety about the risk to life and happiness provoked by the globalization of consumer culture itself. In one study of global brands, the authors found "thirteen percent of consumers are skeptical that transnational companies deliver higher quality goods." They dislike brands that preach American values and don't trust global companies to behave responsibly. Their brand preferences indicate that they try to avoid doing business with transnational firms (Holt, Qulech, and Taylor, 2004, p. 74).

Antiglobalization activism is a diverse movement with some groups promoting various anticorporate or anticonsumption positions. Some call for deconsumption; others dispute specific elements of global value chains such as the use of GMOs (genetically modified organisms) and sweatshop labor; still others promote alternative fair-trade brands; and still others promote local or regional brands such as Mecca Cola (Egypt) or Qibla Cola (Peru), ideologically positioned as explicit competitors to their multinational counterparts (note the influence of global structures of common difference here). Antiglobalization takes many forms, not all of which are truly resistant to consumer culture.

One difficulty of antiglobalization activism lies in the tendency for oppositional forces whether political or cultural to be co-opted by the global mediascape, which increasingly mediates between grass-roots popular cultural manifestations and mainstream corporate consumer culture. For example, Starbucks, the nation of Ethiopia, and the NGO OXFAM waged a recent battle over the right to trademark regional Ethiopian coffee varietal names such as Yirgacheffe, Harrar, and Sidamo. The struggle concerned the right to extract a greater share of the retail value of these regionally branded coffees either for Ethiopian producers or for Starbucks. A vicious court and public-relations battle eventuated in a victory for OXFAM activists and the Ethiopian coffee industry, which may now retain the branding rights (*see* GLOBAL BRANDING: THREE KEYS FOR GLOBAL BRAND SUCCESS). Similarly, in the United States, local coffee houses and regional chains are able to position their brands as anti-Starbucks brands, and incorporate various local, fair-trade, organic, bohemian, and other associations into their brand images that Starbucks cannot claim. But these "doppelganger" brands (Thompson, Rindfleisch, and Arsel, 2006) are as much a part of global consumer culture as the global brands they critique.

Bibliography

Appadurai, A. (1990) Disjuncture and difference in the global culture economy. *Theory, Culture, and Society*, **7**, 295–310.

Cayla, J. and Eckhardt, G.M. (2008) Asian brands and the shaping of a transnational imagined community. *Journal of Consumer Research*, **35**, 216–230.

Clammer, J. (1997) *Contemporary Urban Japan: A Sociology of Consumption*, Blackwell, Oxford.

Dávila, A. (2001) *Latinos, Inc.: The Marketing and Making of a People*, University of California Press, Berkeley.

Featherstone, M. (1991) *Consumer Culture and Postmodernism*, Sage Publications, London and Newbury Park.

Holt, D.B., Quelch, J.A., and Taylor, E.L. (2004) How global brands compete. *Harvard Business Review*, **82**, 68–75.

Latham K. (2006) Introduction: consumption and cultural change in contemporary China, in *Consuming China: Approaches to Cultural Change in Contemporary China*

(eds K. Latham S. Thompson, and J. Klein), Routledge, New York, pp. 1–21.

Lury, C. (1996) *Consumer Culture*, Polity Press, Cambridge.

Mazzarella, W. (2003) *Shoveling Smoke: Advertising and Globalization in Contemporary India*, Duke University Press, Durkham.

Mazzarella, W. (2004) Culture, globalization, mediation. *Annual Review of Anthropology*, **33**, 345–367.

Mintz, S. (2007) Has Chinese Cuisine Survived Six Tumultuous Decades of Communist Rule? The Johns Hopkins Magazine, available at http:// www. jhu. edu/ ~jhumag/ 0607web/ mintz. html; downloaded February 6, 2009.

Sassatelli R. (2006) Virtue, responsibility and consumer choice: framing critical consumerism, in *Consuming Cultures, Global Perspectives: Historical Trajectories, Transnational Exchanges* (eds J. Brewer and F. Trentmann), Berg Publishers, pp. 219–250.

Slater, D. (1997) *Consumer Culture and Modernity*, Polity Press, Cambridge.

Srinivas, A. (2008) *The Indian Consumer: One Billion Myths, One Billion Realities*, John Wiley & Sons, Inc., Singapore.

Stearns, P. (2006) *Consumerism in World History: The Global Transformation of Consumer Desire*, 2nd edn, Routledge, London and New York.

Thompson, C.J., Rindfleisch, A., and Arsel, Z. (2006) Emotional branding and the strategic value of the doppelgänger brand image. *Journal of Marketing*, **70**, 50–64.

Üstüner, T. and Holt, D.B. (2007) Dominated consumer acculturation: the social construction of poor migrant women's consumer identity projects in a turkish squatter community. *Journal of Consumer Research*, **34**, 41–56.

Varman, R. and Belk, R.W. (2008) Weaving a web: subaltern consumers, rising consumer culture, and television. *Marketing Theory*, **8**, 227–252.

Wilk R.R. (1995) Learning to be local in Belize: global structures of common difference, in *Worlds Apart* (ed. D. Miller), Routledge, London and New York, pp. 110–133.

global consumerism and consumption

Russell Belk

A Story of Globalism

In a 1954 *Uncle Scrooge* comic book story by Carl Barks, the rich but penurious Scrooge McDuck is besieged by people wanting favors. A Middle East oil sheik threatens to nationalize his oil wells if Scrooge does not give him a limousine. Others just want money. The pressures of dealing with all this cause Scrooge to have a nervous breakdown. His doctor recommends that he go to a place where there is no money: the land of Tralla La. So Scrooge has one of his pilots fly him, Donald Duck, and the nephews to this mysterious land in the Himalayas. They must parachute into the fog-enshrouded valley, but before they jump Scrooge opens and downs a bottle of the nerve tonic medicine that his doctor has sent with them. The bottle cap falls innocently into the valley and the ducks follow. When they land they find that the natives are blissfully happy and give the ducks meaningful work in their peaceful community. Uncle Scrooge becomes a brick mason and soon finds that his cares and stresses have melted away. Since he no longer has a need for the nerve tonic, he heads to the lake to pour out his remaining six bottles. In the meantime, one of the villagers found the bottle cap that fell from the plane and shows it to his friends. They have never seen anything like it and start a bidding war that quickly escalates into offers of all they have in exchange for the cap. Back at the lake, as Scrooge starts to dispose of his medicine, one of the villagers sees his six shiny bottle caps and excitedly observes that he is the richest duck in all of Tralla La. Scrooge is suddenly back in the onerous position of supreme wealth that he was fleeing. Not wanting to experience the same pressures all over again, he hires Donald and the nephews to trek back to civilization and arrange with his pilots to have 100 plane loads of bottle caps flown in so that he will not be richer than anyone else. After a few weeks the first plane drops its bottle caps on the villagers. They are dancing around in joy when the second and third plane loads descend. Soon they realize that since everyone now has all the bottle caps they want, they are totally worthless. What is more they are beginning to beat down their rice crops, pollute the lake, and cover the ground so that their sheep can find nothing to eat. They turn on the ducks, but Scrooge is able to save them by telling the villagers that the only way to stop the rain of bottle caps is to let them go back to civilization so that they can cancel the order. So with paradise lost and polluted, the ducks return once more to the conditions they have tried to escape.

Although the story was written more than 50 years ago, it can easily be seen as a parable for consumption in an age of globalizing consumerism. The natives of the less affluent world are seen to lust after novel Western goods, at least initially. Later they learn that these things not only do not make them happy, but they also cause them to lose the way of life that formerly brought them joy. Despite his enormous wealth, Scrooge is seen as unhappy with his life until he finds a mythical Shangri-La, which brings him contentment for a time. The bottle caps he brought with him turn out to be the equivalent of money and the source of the evil that he looses upon the valley. In the end, there is no hope for Scrooge to escape the rat race of the more affluent world and, despite his best intentions, he may have ruined the peaceful valley of Tralla La forever. Seen slightly differently, as Dorfman and Mattelart (1984) do in their analysis of the spread of imperialist ideology in Latin America through Donald Duck and Uncle Scrooge comic books, the ducks have brought the seeds of a capitalist exploitative economic system to a blissful communist society. But because the ducks leave Tralla La, the reader does not get to learn whether the outcome of this economic imperialism is ultimately good or bad.

For a glimpse at possible outcomes of unleashing Western-style consumerism on other parts of the world, two films invoking Coca Cola are relevant: Uys' (1983) *The Gods Must be Crazy* and Norbu's (2000) *The Cup*. In *The Gods Must be Crazy* it is a Coke bottle that falls from an airplane and it is found by !Xi, a !Kung bushman of the Kalahari Desert in Botswana. When he shows the strange bottle to others in his hunter gatherer society, they initially try using the bottle in their daily tasks of preparing food and producing bark cloth. But eventually the iconic bottle becomes,

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

like the bottle caps in Tralla La, an object of intense envy, lust, hatred, and violence. The sharing-based subsistence economy has been spoiled by the introduction of a nonindigenous good along with the implications of private property, materialism, greed, and the global consumer culture from which the bottle derives (*see* CONSUMER MATERIALISM). The only way to dispel this snake from the Eden of the Kalahari is for !Xi to take the bottle to the edge of the earth and throw it off.

The Cup is a rather different film about the introduction of Western things (a crumpled Coke can and television) into an unspoiled society (a Tibetan Buddhist refugee monastery in the Himalayas, not unlike Tralla La). The young monks in the film love soccer and as the film begins, they are playing the game using a crumpled Coke can as a ball. Coca Cola's iconic script is shown as a counterpoint to the prayers carved into mani stones and inscribed on the prayer wheels of the monks. The young soccer players scatter when a stern senior monk, Geko, appears. Geko picks up the can and brings it to the Lama who already has several reclaimed Coke cans on his altar serving as lamps burning clarified yak butter as a part of Tibetan Buddhist worship and prayers for the dead. Here, rather than the Western consumption object unleashing the ills of Western civilization on an innocent group of people, the foreign object is domesticated, localized, and incorporated into sacred rites. Just as the monks later accept soccer and satellite television, they have found a way to transform the profane, commercial, foreign object into a sacred, traditional, local good that accentuates rather than disrupts their spiritual lives.

Each of these stories presents a different take on the local consequences of global consumerism. We might well ask whether the outcome of globalizing consumerism is closer to the benign assimilationist scenario of the Tibetan Monastery or whether it is, instead, one of the more malignant disruptive versions seen in Tralla La or the Kalahari. What follows is a brief attempt to answer this question on the basis of theory and evidence.

Global Consumerism

Global trade circuits like the Silk Road and the Spice Route are ancient and suggest that consumer goods have moved across cultures for some time. But several features make contemporary globalization different. The speed of transportation and technological developments mean, for example, that fresh strawberries can be had in most places in the world at any time of the year. More goods travel greater distances more quickly than ever before. Many of these goods are branded products and the multinational companies that market them make special efforts to promote them in local languages and media. Media has also globalized, however, and film, television, radio, Internet sites, telephone connections, magazines, news programs, and sporting events can now often have global reach (*see* GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES). These connections mean that we have windows on the world's consumption like never before. Japanese manga and American rap music not only cross the boundaries of their countries of origin, they also inspire local variations like American graphic novels and Japanese hip-hop culture (*see* GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES). Beyond the specific brands and program content that are globalizing and cross-fertilizing one another, the basic consumption ethos and the place of consumption in everyday life of people are changing as a result. Accordingly, we find Western-style shopping malls in Egypt, Christmas gift-giving celebrations in China, and Bollywood movies in Qatar.

One gloss for such developments is the advance of global consumerism. Consumerism, or consumer culture (*see* CROSS-CULTURAL PSYCHOLOGY OF CONSUMER BEHAVIOR), has been analyzed as having four defining characteristics:

1. A substantial portion of the population consumes at a level substantially above subsistence.
2. Exchange dominates self-production.
3. Consuming is accepted as an appropriate and desirable activity.

4. People judge others in terms of their consuming lifestyles (Rassuli and Hollander, 1986).

However, the first of these criteria may be too restrictive. Even for the poor of the world, the allure of consumer goods can sometimes be so strong that they are willing to sacrifice "necessities" in order to be able to afford "luxuries." Examples include Romanians sacrificing eating in order to afford a refrigerator into which they cannot afford to put food, and Turks selling the milk-producing family cow to be able to buy candy for their children (Ger and Belk, 1996). Such practices have been referred to as *leaping luxuries* (Belk, 1999). Rather than simply see things as necessities or luxuries, a third transitional category is that of decencies – things like deodorants, feminine hygiene products, soaps, and hot running water. And as global communications promote world standards in such products, they also move from being seen as decencies to being regarded as necessities. Stearns (2002) suggests another criterion for a consumer culture to exist: deriving part of our identity from consumption objects; in other words, embracing a notion of the self extended through possessions. Together, at both the cultural and the individual levels, these criteria suggest a heightened consumption orientation in our lives.

By focusing on the broad attitudinal and value-based concept of consumerism rather than the mere fact of the proliferation of global brands and multinational corporations, we can see that the presence of global brands may or may not signal global consumerism. It also means that not only products but services also can be a part of this consumption orientation. And it means that the celebration of globalizing consumption holidays like Christmas may also tell us something about the interpenetration of global and local meaning systems. One possibility is that participation in global holidays provides a sense of shared global identity. Nevertheless, Belk and Kimura (2005) find that Christmas in Japan is localized and involves couples going out rather than families remaining in or returning to the home. Similarly, Yoko Brannen (1992) shows how Tokyo Disneyland has been adapted, domesticated, and made to fit within Japanese culture, despite an outward insistence by the owners that the theme park be an exact replica of the one in California. This is done by both making the exotic park familiar (e.g., employees use the polite Japanese service voice rather than the more brash approach of Anaheim's Disneyland), and keeping the exotic exotic (e.g., in park performances American, Chinese, and Korean characters are caricatured and made to seem totally other, even though Japan has undeniably been influenced by each of these cultures). Further examples can be found in the adaptations that McDonald's has had to make for consumers in Asia (Watson, 1997). Rather than imposing a universal formula, the chain has had to alter menus, sacrifice quick table turnover, and realize that the restaurant can be seen by some as a significant fine dining experience reserved for special occasions and not simply a convenient place to get a meal. Global media have also had to adapt to local cultures. MTV had attempted to use a globalized approach in its Asian programming until Channel V began to localize for Indian audiences. It was so successful, that MTV was forced to do the same.

Discussions and theories concerning the effects of global consumerism have gone through several phases. The first involved theories of cultural imperialism as Western brands, entertainment, and media were seen as being imposed on the rest of the world (Dorfman and Mattelart, 1984). As we have seen, this thesis has not proven tenable, in part because consumers are active rather than passive recipients of marketing offerings. A second set of theories saw the reaction to global products, services, and media as producing a split between the West and the Rest. A variation of this theme sees the clash as being between the affluent minority and the poor majority. A third wave of theorizing saw the local as resilient and powerful enough to coopt, reconfigure, or reinterpret the global for its own purposes. A variant of this theme sees local businesses as able to use their cultural knowledge to out-compete global brand adversaries. If the cultural imperialist position erred on the side of seeing the consumer as too passive, resilient, and powerful, the local position likely sees the consumer as too active and empowered. A fourth perspective sees the globalization of consumption as fracturing into

a prism of global, local, and hybrid offerings with the latter creolized or glocalized fusions becoming the most common (Appadurai, 1996; Ger and Belk, 1996). Appadurai (1996), for example, focuses on the impacts of global flows of people, ideas, finances, media, and technology and the resulting disjunctures this creates in economies and cultures. Ger and Belk (1996) add to this list the changes and instability brought about by global flows of consumption objects that accompany these other mobilities. Since cultures constantly change and interact with each other, it seems reasonable to expect that certain intercultural elements will blend with one another in unique ways as they do so. Thus Wal-Mart, Ikea, and Carrefour are not so much spreading American, Swedish, and French culture throughout the world as they are mediating, combining, and reshaping cultures both in their home countries and abroad. A variation on this theme of dynamic hybridity is Wilk's (1995) notion of global structures of common difference. Wilk (1995) uses the example of the global spread of beauty pageants to suggest that, rather than resulting in any homogenizing outcomes, it is the structure of such competitions that is being globalized. In other words, the beauty contest and its focus on certain kinds of differences is diffusing globally, but this does not result in a hegemonic global standard of beauty. Global beauty contests instead spread the focus on structures of youth, race, ethnicity, and power guiding theses contests.

Consumption of the Global

One alternative to embracing global consumerism is the rise of nationalism (Ger and Belk, 1996). This was the case, for example, in Chinese efforts to resist the flood of global goods in Old Shanghai before communism. There is some evidence of a rise of pride in Chinese brands in contemporary China as well. And there is some evidence of contemporary Indian feelings of nationalism in resisting foreign brands and stores (Varman and Belk, forthcoming). Eventually, the distinction between global and local brands and stores may disappear. Watson (1997) tells the story of Japanese Boy Scouts visiting Chicago and discovering that America also has McDonald's; and how, in Hong Kong too, most young people are unaware of the company's American origins. One of the reasons for this decline in recognition of globalism is simply habituation. First generations may recognize the entry of the global, but for subsequent generations, these origins are outside of their awareness. And even those Russians who are aware of the foreign origins of products like Coca Cola and McDonald's hamburgers may simply take them for granted and cease to think about their origins. A second reason for not attending the global is that with local franchises, joint ventures, and outsourcing, the distinction between what is local and what is global is blurring. Are Toyotas made in Canada any more foreign than Blackberry devices made in China? When companies on different continents merge, it becomes even harder to distinguish the local brand. At a more general level, what we in the West know as Chinese food is often not recognized as "real" Chinese food by Chinese visitors or immigrants. And as a part of a dynamic culture with various regional cuisines often fusing, Chinese food in China is also a changing category.

One of the reasons for the initial concern with resisting "cultural imperialism" in globalization and subsequent championing of the local was the feeling that something authentic is lost when we begin to replace local products and services with global goods. In Ritzer's (2004) view, the unique one-of-a-kind product is replaced by the interchangeable, dehumanized, and disenchanted commodity, or what he calls *nothingness*. There is both a romanticism and an elitism in this view. But to jump to the other extreme and claim that globalization is good for everyone would also be shortsighted. As with the opening of Eastern Europe after the fall of the Berlin Wall, one of the initial consequences of a flood of novel goods may be increased materialism. There is evidence that materialism is negatively related to happiness and feelings of well-being. Nevertheless, the direction of causality in this relationship is not entirely clear and Schudson (1998) argues that the desire for material goods may bring dignity and rationality to peoples' lives. But even if this is true for some people, those shut out of consumer culture by poverty may end up feeling even more deprived by

upsurges in global consumer culture. And at a more macro level, high levels of consumption fueled by globalism are likely unsustainable and harmful to the environment.

In the end, rather than seeking a Shangri-La free from global capitalism and global money or feeling that our way of life has been corrupted by the sudden arrival of a Coke bottle, we will very likely come to simply accept global production and consumption as the incontrovertible way of the world. Using Appadurai's (1996) notion of global flows, we can think metaphorically in terms of that most liquid of flowing substances, water. We are unlikely to view the water in our rivers and seas as being either local or foreign. It comes from somewhere and goes somewhere without bearing any traces of its origins or destinations. Granted when an iceberg shows up off the coast of southeastern Canada, it is recognized as coming from somewhere else. But the longer it is there, and the more it melts away, the less foreign it seems. And as for those bottles of Perrier, their claim of foreign origin is as much a matter of image as it is ownership. Perrier is now owned by the Swiss conglomerate Nestle, in any case. Other companies like Louis Vuitton remain French, but their products are no longer made in France. None of this is to say that globalism is inconsequential. The upsurge of globalism starting in the late twentieth century has resulted in some of the most sudden and profound economic shifts in the history of the world. The economic rise of China and other Asian economies are a dramatic evidence of these shifts. But from the perspective of Western consumerism and consumption, most of these changes were relatively invisible until the recalls of a number Chinese-made products in 2007.

A far greater and more sudden result of China's economic rise in the past 30 years is what is widely described as the country's *consumer revolution* (Davis, 2000). Part of this revolutionary rise of consumer culture has been the influx of foreign brands, stores, and advertising (Zhou and Belk, 2006). But much of it is based on rapidly rising incomes and a big push from the Chinese government to develop domestic consumption in order to reduce dependence on exports (Croll, 2006). However, not everyone has benefited from these changes. Worker demonstrations have been common as state-run enterprises close down causing unemployment and breaking the once solid iron rice bowl of guaranteed employment. The cities, and especially the southern coastal cities of China, have benefited far more than rural areas and China's west. Even so, there remain many urban poor, and families in rural areas experience social upheaval when one parent goes off to work in the city for extended periods. The young are better able to adapt to the new economy than the old, and there is resentment between the new haves and the new have-nots of China (Croll, 2006). Even the "little emperor" children who are blessed with indulgence by parents and grandparents thanks to China's one-child policy, suffer as they become more obese and feel the added pressure of carrying the family's face. Few of these problems are unique to China, but they are made prominent by the rapidity of changes in Chinese consumption and the contrast to the days of Mao, the Cultural Revolution, and the Great Leap Forward. Chinese consumption bears close watching because it makes the effects of consumer culture highly visible as it shrinks the time period over which it unfolds and distills the best and the worst of global consumerism.

Bibliography

Appadurai, A. (1996) *Modernity at Large: Cultural Dimensions of Globalization*, University of Minnesota Press, Minneapolis.

Belk, R. (1999) Leaping luxuries and transitional consumers, in *Marketing Issues in Transitional Economies* (ed. R. Batra), Kluwer, Norwell, pp. 38–54.

Belk, R. and Kimura, J. (2005) Christmas in Japan: globalization versus localization. *Consumption, Markets and Culture*, **8**, 325–338.

Croll, E. (2006) *China's New Consumers: Social Development and Domestic Demand*, Routledge, London.

Davis, D.S. (2000) *The Consumer Revolution in Urban China*, University of California Press, Berkeley.

Dorfman, A. and Mattelart, A. (1984) [original 1971] *How to Read Donald Duck: Imperialist Ideology in the Disney Comic*, International General, New York.

Ger, G. and Belk, R. (1996) I'd like to buy the world a coke: consumptionscapes of the 'less affluent world'. *Journal of Consumer Policy*, **19** (3), 271–304.

Norbu, K. (2000) The Cup, 94 minute film, New Line Home Video.

Rassuli, K.M. and Hollander, S.C. (1986) Desire – induced, innate, insatiable? *Journal of Macromarketing*, **6** (Fall), 4–24.

Ritzer, G. (2004) *The Globalization of Nothingness*, Pine Forge Press, Thousand Oaks.

Schudson, M. (1998) Delectable materialism: second thoughts on consumer culture, in *Ethics of Consumption: The Good Life, Justice, and Global Stewartship* (eds D.A. Crocker and T. Linden), Rowan and Littlefield, Ann Annapolis, pp. 249–268.

Stearns, P. (2002) *Consumerism in World History: The Global Transformation of Desire*, Routledge, London.

Uys, J. (1983) The Gods Must be Crazy, 109 minute film, 20th Century Fox.

Varman, R. and Belk, R. (forthcoming) Nationalism and ideology in an anti-consumption movement, *Journal of Consumer Research*.

Watson, J. (ed.) (1997) *Golden Arches East: McDonald's in East Asia*, Stanford University Press, Stanford.

Wilk, R. (1995) Learning to be local in Belize: global systems of common difference, in *Worlds Apart: Modernity Through the Prism of the Local* (ed. D. Miller), Routledge, London, pp. 110–133.

Yoko Brannen, M. (1992) 'Bwana Mickey': constructing cultural consumption at Tokyo Disneyland, in *Re-Made in Japan: Everyday Life and Consumer Taste in a Changing Society* (ed. J. Tobin), Yale University Press, New Haven, pp. 216–235.

Zhou, N. and Belk, R. (2006) Chinese consumer readings of global and local advertising appeals. *Journal of Advertising*, **33** (4), 63–76.

strategic export marketing – achieving success in a harsh environment

Constantine S. Katsikeas and Leonidas C. Leonidou

Introduction

The modern business environment is characterized by accelerating globalization, intensifying competition, and growing integration, which have been responsible for inducing an ever-increasing number of firms to embark on international operations (Kotabe and Helsen, 2001). As opposed to other more direct modes of foreign market entry such as franchising (*see* INTERNATIONAL FRANCHISING), joint venturing, and wholly owned production, exporting has been more widely adopted by manufacturing firms because it is less risky, more flexible, and less costly Albaum and Duerr, 2008. In fact, exporting is the most popular, the quickest, and the easiest way for a company to become international, because it offers numerous benefits: it improves the firm's financial position, helps to transfer technologies and know-how from other countries, enriches managerial skills through exposure to idiosyncratic foreign environments, makes better use of production capacity, and facilitates organizational growth (Czinkota and Ronkainen, 2006).

The deployment of successful export marketing strategies constitutes a critical, and in many cases, a thorny process, which could assist in reaping the above benefits while at the same time avoiding the dangers involved (Leonidou, Katsikeas, and Samiee, 2002; Katsikeas, Morgan, and Kaleka, 2004). To be effective, such strategies should rely on a thorough analysis of the current situation, particularly focusing on the firm's resources and capabilities, the export task environment, and the foreign macroenvironment. It will then have to carefully target export markets through systematic foreign market selection, as well as through customer targeting and positioning. Following this, it is important to set achievable objectives that will enhance the firm's effectiveness, efficiency, and adaptiveness. Strategies for each of the elements of the marketing mix should subsequently be designed, with special focus on the right degree of adaptation/standardization (*see* STANDARDIZATION/ADAPTATION OF INTERNATIONAL MARKETING STRATEGY). Finally, certain issues for organizing and controlling the exporting strategy should be examined.

Export Situation Analysis

As a starting point, the firm first needs to evaluate the existing situation with regard to its internal resources and capabilities, the export task environment, and the foreign business environment.

Company resources and capabilities. The firm's resources and capabilities are important prerequisites for achieving sound export marketing strategies (Leonidou, 1998; Leonidou, Katsikeas, and Piercy, 1998). Their analysis will reveal the strengths that are essential to support its exporting effort, as well as any weaknesses to be removed. With regard to *resources*, a major driving force is the specific interest and commitment of top management to exporting, particularly as regards the belief that export engagement is beneficial to the firm. Previous experience of managers in foreign countries and their special international marketing skills are two additional factors that safeguard the proper design of export strategies. Research and development represents a vital resource for exporting, especially when a firm has proprietary technical knowledge for unique products. The same is also true for production resources, such as the existence of modern manufacturing technology and the availability of adequate production capacity. Information resources are also crucial in building sound export strategies, especially if they refer to overseas market demand, foreign business practices, and export regulations and documentation. Finally, adequate funds to finance the export activity, coupled with the existence of a specialized export department to deal with international business, are also important to support export strategies.

On the *capabilities* side, it is very important for the firm to be able to identify business opportunities, such as locating and analyzing the potential of foreign markets, contacting prospective customers abroad, and acquiring specialized information for specific international projects.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Another key capability is that of relationship building, particularly obtaining reliable representation in overseas markets and subsequently establishing and maintaining close relationships with foreign customers. Product-market development capabilities are also essential for building successful export strategies, since international markets provide a fertile ground for the adoption of new production methods, the development of innovative products, and the application of novel marketing techniques. A final point refers to the firm's capability to adapt to environmental changes, because the diverse and volatile nature of the foreign business environment requires a certain adjustment in the elements of the marketing mix.

Export task environment. The analysis of the export task environment will provide the exporting firm with crucial knowledge about the market, the buyers, the competition, and the marketing intermediaries (Leonidou, Kaminarides, and Panayides, 2007). These are forces directly affecting the firm's export marketing activities.

With regard to the *market*, the firm first needs to estimate the size of the market (both in terms of volume and value) in various foreign countries and determine its growth. It is also vital to define the structure of the market, since this will guide the firm to formulate its pricing strategies in the right context. Some other key issues relate to demand differences, demand seasonality, and demand elasticity, caused by variations in income levels, product images, and brand loyalty. Any barriers associated with entering (e.g., long geographic distances, trade barriers, and government regulations) and/or exiting (e.g., nonrecoverable investments, binding agreements with distributors, and personnel hired for a long period of time) foreign markets are also to be examined.

Apart from general foreign market characteristics, it is important to analyze the characteristics of foreign *buyers*. The habits and attitudes of buyers are not identical around the world, because of variations in topographic and climatic conditions, household size and structure, level of technical understanding, income level and its distribution, educational standards, manners and customs, and so on. Consumer buying roles (e.g., decider, purchaser, user) also need to be examined since they differ across countries, and the same is true with regard to the purchasing power, shopping behavior, and usage patterns of end users. This is because all these lead to different product preferences, price acceptance levels, distribution systems, logistical support, and communication methods.

Competition is one of the most crucial components of the export task environment and, as such, it warrants special attention. Although the firm may enjoy a competitive advantage in the domestic market, when transcending national boundaries it may lose this advantage owing to more complicated, multiple, and intense competitive situations. The exporter should focus on the following issues relating to competition in international markets: origin of competitors (e.g., host country, home country, multinational), source of competition (e.g., same product market, related technology, substitutes), competing base (e.g., low cost, product differentiation, government protection), market position (e.g., leader, challenger, follower), and major strengths/weaknesses (e.g., financial, managerial, product).

The last component of the task environment concerns foreign *marketing intermediaries* (*see* INTERNATIONAL MARKETING CHANNELS), who help the company promote, sell, and distribute its products abroad. First, the exporter has to examine the structure and operation of the distribution system, since this implies adjustments in the way a product can be made available to the end user. Such adjustments should take into consideration the major actors in distribution, and the services they provide. The peculiarities involved in communicating effectively with end users in different countries necessitate an investigation into the availability of advertising agencies. Equally, it is useful to identify any research agencies that will help to provide timely, reliable, and accurate information about foreign market characteristics. Finally, the cost and quality of other types of marketing intermediaries, such as transportation/shipping firms, financial institutions, and mass media organizations, also need examination.

Foreign macroenvironment. The foreign macroenvironment consists of forces that are broad in nature, usually uncontrollable, and the source of opportunities to be exploited and threats to be avoided. However, as opposed to the domestic environment, the foreign environment is characterized by greater complexity, uncertainty, and heterogeneity attributable to the many different countries that comprise the international market. In evaluating the foreign macroenvironment, it is important to analyze its physical, demographic, economic, sociocultural, political–legal, and technological aspects (Leonidou, Kaminarides, and Panayides, 2007).

With regard to the *physical* environment, large geographic distances between the home and foreign markets and different territorial sizes among countries seriously influence transportation costs, product availability, pre- and postsales service, and product quality in the foreign market. Differences in climatic conditions also shape consumption patterns in each country and determine product/packaging requirements. The foreign country's natural resources are important in order to determine its wealth, as well as explore the potential to source raw materials for the company. The growing trend in many countries toward protecting the natural environment exerts pressure on the exporting firm to produce eco-friendly goods.

In the *demographic* environment, it is important to obtain information on the population size and growth of foreign countries, since these are key indicators of market attractiveness. It is also useful to know the age structure of the foreign country's population, as it is widely understood that consumers have different needs at various stages of their lives. The way the population of a country is geographically distributed largely affects the effectiveness and efficiency of certain marketing activities, particularly those relating to distribution, logistics, and promotion. Another demographic element that affects consumer behavior is the social class stratification of the population.

The *economic* environment consists of factors that affect consumer purchasing power, spending patterns, and living standards in overseas markets. Here, it is important to study per capita income and the way it is distributed in each country, since this is inextricably linked to the size, quality, and potential of a foreign market. Foreign exchange rates and controls should also be examined largely, because they critically affect demand for the company's products abroad. Pricing strategies are greatly influenced by inflation rates (and price controls), which should be clearly identified in each country. The foreign economic infrastructure (e.g., road network, ports/airports, energy) also needs to be investigated, because it influences distribution methods, logistics costs, and promotional activity.

Understanding the *sociocultural* environment is crucial because it affects the basic values, perceptions, preferences, and behavior of the society. Differences in religion, values and attitudes, manners and customs, aesthetics, education, and social organization can greatly affect consumer behavior, targeting approaches, and marketing programs. This situation becomes more complicated in the case of countries with many subcultures and/or different cultural contexts. The exporter should also become familiar with the oral, written, and nonverbal aspects of the language used in a specific country, because these affect branding, labeling, and advertising decisions. Lack of familiarity with foreign business practices may place the exporter in a stressful situation, particularly as regards to building relationships with overseas distributors. For instance, while personal rapport is vital in conducting business in some countries, in others more formal procedures have to be followed.

The *political–legal* environment consists of government agencies, laws, agreements, and pressure groups that can seriously influence the firm's export operations. Foreign governments may impose a number of controls on exporters, such as entry barriers, price restrictions, and exchange controls that can turn the exploitation of export opportunities into a tedious, expensive, and prolonged task. Import tariffs may also pose a serious problem for the exporting firm, since they cause export prices to escalate, while nontariff measures (e.g., quantitative restrictions) may restrict the company's potential in overseas markets. Some of these markets are plagued by political instability caused by economic, societal,

and/or political factors, which may seriously reduce strategic marketing effectiveness.

The last component of the foreign macroenvironment refers to *technology*-related issues. A crucial point here is the level of technology used by buyers in the host country. This will affect product design, after-sales service, and communication modes. Understanding the prevailing communications technology is essential in order to achieve effective interaction with foreign partners, collect useful information from overseas markets, and design sound advertising strategies. Transportation technology is also crucial, because it affects the safe and speedy delivery of the products. Finally, the exporting firm needs to explore the degree to which information technology can be used, especially the Internet, because it has such a strong influence on the way the firm can communicate and sell its goods to middlemen and end users.

Export Market Targeting

Following the analysis of the internal and external situation, the firm has to select the most attractive foreign countries to enter, as well as define the target customer groups to supply to within each country (Majaro, 1982).

Foreign country selection. The exporting firm has to analyze various countries in order to *select* those with the greatest appeal. To this end, world markets should be systematically *segmented* based on a number of "macro" criteria, such as geographic location, market size, growth potential, cultural characteristics, stage in product life cycle, level of competition, and host government attitudes. This broad segmentation of international markets is essential to identify foreign country markets that could be easily accessible, approached in an effective manner, and yield adequate financial returns.

Having selected the countries that are most attractive as export destinations, the next step is to define the number of countries to be approached. Here, the firm has to decide about its *export market expansion* strategy, that is, whether to concentrate on only a few countries or to extend its activities to multiple countries (Katsikeas and Leonidou, 1996). While the former may help to better understand specific customer requirements, establish and maintain long-term market presence, and maximize returns from individual markets, the latter can enhance product specialization, spread business risks, and result in greater flexibility. Since both strategies offer certain advantages (and disadvantages) for the exporter, the choice should depend on a number of contingent factors pertaining to the company (e.g., risk perceptions), the product (e.g., stage in life cycle), and the market (e.g., market stability).

The exporter has then to *prioritize* the countries it has selected to serve, and set the sequence and speed of entering each of them. In doing so, it has to take into consideration the availability of internal resources (e.g., financial, manpower, know-how), the attractiveness of the country concerned (e.g., market potential, level of competition, buyer sophistication), and the conditions of the global business environment (e.g., economic indicators, regulatory framework, technological change). It is also important to specify whether the firm should enter the market indirectly (i.e., selling through independent organizations located in the home country) or directly (i.e., selling directly to an importer or buyer located abroad). As opposed to indirect exporting, which is a low-cost and less complicated method of engaging in international sales activities, direct exporting allows for greater control and better exploitation of international markets.

Country target groups. Within each country selected, it is important to identify those market segments that are the most appropriate. Here, the exporter needs to *segment each foreign market* according to a number of geographic (e.g., region), demographic (e.g., age group), psychographic (e.g., life style), behavioral (e.g., product usage), or other bases that are considered particularly suited to a firm's products. This exercise will reveal various segments that should be evaluated on a number of criteria, such as size/growth pattern, intensity of competition, and firm resources. It is essential to note that target segments may vary across countries, due to the existence of different stages of market development.

Having identified the segments within each foreign market, the exporter has then to *select*

the number of and *set the sequence* in which the segments need to be served. Here, the firm may either concentrate on a specific market segment (or segments) in a country using the same marketing mix, or adopt a differentiation strategy where each target segment is approached using a different marketing mix. The selection of the specific strategy will depend on a number of factors, namely availability of company resources, nature of the product, stage of product life cycle, degree of market variability, and competitors' market coverage strategies.

Finally, the exporter has to establish the *positioning* of its products in the minds of consumers in relation to competing brands. To succeed, it is important first to arrive at points of parity with the competition, identify areas with a competitive advantage, and choose those that are most promising to decide on the overall positioning concept. Most importantly, the firm needs to build its marketing mix around this concept, maintain the positioning chosen over time, and monitor market changes so as to make the necessary adjustments. Because of the diversity of overseas countries, the firm has to decide on whether to use the same positioning theme worldwide or tailor this to each individual market. While the former is more suitable in the case of catering to universal segments worldwide, the latter is more applicable when segments encounter cultural differences, legal constraints, and competitive forces across countries.

Export Objectives

Having analyzed the current situation and decided on the targeting and positioning strategy, the exporter needs to set clear, measurable, and achievable objectives that will provide direction to the exporting strategy (Hibbert, 1985). First, the firm needs to establish its *overall export objectives*, such as export sales volume, export profits, and return on investment from export activities. Then, these objectives have to be broken down for each of the target countries selected, taking into consideration demand conditions, intensity of competition, and other market peculiarities. Finally, for each target segment within each country, it is important to determine specific objectives, such as market share, sales, and profitability. Notably, the exporter should set not only financial but also nonfinancial objectives, such as those relating to the satisfaction, attraction, and retaining of customers.

Overall objectives should correspond to specific *marketing objectives* for each target market/segment, namely, those relating to product (e.g., new-product introductions), price (e.g., profit margins), distribution (e.g., distribution coverage), logistics (e.g., product delivery time), and promotion (e.g., brand awareness levels). In addition, the exporter should distinguish between strategic (e.g., deciding to use distributors in some countries, as opposed to agents in others) and tactical (e.g., selecting a particular distributor or agent within a specific country) objectives.

Export Marketing Mix

To achieve the firm's export objectives, a sound strategy has to be designed for each of the elements of the marketing mix (i.e., product, price, distribution, logistics, and promotion). Apart from achieving consistency, synergy, and synchronization among these elements, a key issue that needs to be addressed here is the degree of standardizing or adapting the marketing strategy in each of the targeted overseas markets (Cavusgil, Zou, and Naidu, 1993; Cavusgil and Zou, 1994). However, the decision to standardize or adapt the marketing strategy is situation specific, and this should be the outcome of a thorough analysis and assessment of the relevant contingency factors prevailing in a specific market at a specific time. In addition, the appropriateness of the selected level of strategy standardization/adaptation should be evaluated on the basis of its impact on company performance in export markets (*see* STANDARDIZATION/ADAPTATION OF INTERNATIONAL MARKETING STRATEGY).

Product. Different conditions of use, variations in purchasing power, dissimilar consumer tastes, and diverse sociocultural settings imply adjustments of the company's product *design and ingredients* to meet the idiosyncrasies of each foreign market. Moreover, in some countries, consumers are very *quality* conscious, thus requiring further improvements of the product.

Attention should also be paid to *branding* issues (*see* GLOBAL BRANDING: THREE KEYS FOR GLOBAL BRAND SUCCESS), since some brand names are difficult to pronounce and/or have some undesirable meaning or associations in certain languages. The special conditions involved in selling the product abroad (e.g., large geographic distances, rough product handling, extensive shelf time), coupled with the stringent safety legislation prevailing in some countries, requires more protective and safer *packaging*. Package appearance and size may also need adjustment, owing to variations in consumer preferences, disposable incomes, and shopping frequency.

With regard to *labeling*, legislation in most countries requires the provision of full information regarding product ingredients, nutritional value, production/expiry dates, and other elements, which have to be expressed in the local language(s). Certain symbols, signs, and colors appearing on the label may have to be changed to conform to local traditions, religion, and values. Differences in product use, availability of after-sales service, and competitive practices across countries necessitate changes in the nature and duration of the *warranties* offered for the goods exported. Finally, the provision of an *after-sales service* has to be adjusted in certain countries owing to different product usage, unavailability of spare parts, and poor capabilities of local firms to provide such a service.

Price. Offering satisfactory *prices* is key to attracting export customers, especially in light of the fierce competition that characterizes international markets. However, as opposed to domestic prices, export prices usually escalate because of additional costs incurred in modifying the product, its packaging, and service in overseas markets; higher administrative, operational, and transportation costs connected with exporting; extra taxes, tariffs, and fees imposed when entering the host country; and the higher cost of marketing and distributing goods in foreign markets. Certain unexpected factors, such as unfavorable foreign exchange rates, may make the company's prices even less competitive. Reduced competitiveness can also occur because of the higher cost of distribution, logistics, and promotion abroad, adoption of dumping practices by overseas competitors, and foreign government policy to subsidize the local industry.

The exporter should also be cautious in quoting *sales terms* (e.g., cost, insurance, and freight) that are mostly preferred by foreign customers, as well as *credit terms* (e.g., 90 days letter of credit) that the industry in which it operates is accustomed to offering. Both sales and credit terms involve certain risks, which tend to be greater for foreign customers, who are much farther away, have no past experience with the company, and come from countries with unstable politico-economic environments. Another key aspect refers to the *currency* in which the export quotation is made, which has to be less vulnerable to foreign exchange fluctuations, and contribute toward making the company's prices attractive to overseas buyers.

Distribution. Obtaining *reliable representation* in foreign markets is one of the major challenges for achieving export success. It is very difficult to find foreign representatives who meet the structural (e.g., territorial coverage, financial strength, physical facilities), operational (e.g., product assortment, logistical arrangements, warehouse facilities), and behavioral (e.g., market reputation, relationships with government, cooperative attitude) requirements of the exporting firm. Once representation is secured, it is important to establish mechanisms (e.g., personal visits abroad, electronic communication, provision of financial incentives) to maintain control over middlemen abroad. Such mechanisms are imperative in export markets since there is greater dependence on the middleman due to binding legal agreements; it is difficult to find other middlemen in the overseas market as replacements; and the middleman carries other product lines that may be more profitable than those handled by the exporter.

On entering foreign markets, the exporter is usually faced with *distribution systems* that are different from those in the home market. Distribution differences are also likely to exist among international markets themselves. For example, as opposed to developed countries, developing nations are characterized by a higher per capita

number and smaller-sized retail outlets. Moreover, while in some countries, distribution channels consist of many layers, in others, direct distribution systems are more evident. Furthermore, the range and quality of the services offered by channel members vary considerably across countries. These variations imply that the firm's *distribution methods* have to be adapted according to the idiosyncrasies of each foreign market. The exporter should realize that, in some countries, access to distribution channels is not always available because they are already occupied by the competition, the length of the channel may be too costly to manage, or the power may rest with a certain distributor who controls entry at various levels of the system. To bypass this problem, the exporter can piggyback on an already established distribution mode by another exporter who sells complementary goods, seek the assistance of export management companies, or set up its own distribution channels.

Logistics. Logistics (*see* DESIGNING A GLOBAL SUPPLY CHAIN: OPPORTUNITIES AND CHALLENGES) is the element of the export marketing mix that requires immense adaptation, owing to the fact that products have to cross national boundaries, the multiple markets that the firm has to supply, and the peculiarities that these markets have with regard to their infrastructure. Beginning with the *ordering procedures*, the firm has to become familiar with the specific export documentation, shipping arrangements, and customs procedures required for sending its goods abroad. Ordering procedures for exports are generally more complicated than those for domestic sales, because they demand numerous documents, involve many parties, and incur high costs. Hence, it is important to have people within the company specializing in these procedures and/or to seek the advice of specialists in the field, such as foreign freight forwarders. The ultimate aim is to secure smooth order taking and execution procedures to satisfy foreign customers.

With regard to *transportation*, distances from foreign markets are usually greater than is the case domestically, thus involving greater risk of damaging the product, delaying product delivery, and increasing transportation costs. This situation is even more critical in countries with large interborder distances, poor infrastructural facilities, and limited availability of transportation means. All the above factors increase transportation, insurance, and other costs associated with exporting, which are subsequently passed on to final selling prices. The exporter should be aware of the alternative modes of transporting goods abroad, as well as securing the best possible terms.

Finding adequate *warehousing* facilities in the host market is critical in ensuring a constant product flow, achieving timely delivery, and maintaining product quality at high levels. However, in some countries there are neither warehouses available to store the company's products nor proper installations to safeguard their quality. Storage fees may also be prohibitive in some countries, while in others warehousing equipment technology may be outdated. Storage problems become more acute in the case of countries where there is a large territory to cover and the consequent need for a multiple warehousing system to obtain satisfactory market coverage and support. The exporter can circumvent these problems by finding distributors with adequate warehousing space and technology and, if this is not feasible, by setting up its own warehouses.

The firm should handle its *inventory* policy with care in overseas markets, because transportation delays, demand fluctuations, and unexpected events can create shortages of its products. This is usually more critical when dealing with geographically distant countries, where it takes more time to resupply the market. Such a situation may involve a number of disadvantages for the exporter, such as (i) noncredible image of its products in the foreign market; (ii) lost sales and profits from potential and existing customers; and (iii) extra costs in trying to replace product shortages by using faster and more expensive transportation means. To accommodate this situation, the exporter has to put in place a sound inventory control system, which should take into consideration the territorial size, infrastructural facilities, and purchasing/consumption habits of each foreign market.

Promotion. Variations in buying motives, consumption patterns, and government regulations make it necessary for the firm to adjust its promotional activity to fit individual foreign market requirements. Of the elements of the promotion mix, *advertising* (*see* INTERNATIONAL ADVERTISING – IS THERE STILL A STANDARDIZATION VERSUS LOCAL ADAPTATION DEBATE?) requires particular attention, not only because it habitually takes the lion's share of the promotional budget but also because of its greater sensitivity to the above differences. There are several specific issues in international advertising that need attention, namely, variations in the composition of the target audience; inappropriate content of the advertising message; unavailability or different use of advertising media; restrictions in the frequency/duration of advertising; and insufficient means of assessing advertising effectiveness. These factors are responsible for making adjustments in the advertising message, its execution style, and the media mix used across countries.

What is an effective *sales promotion* in the firm's home country is not necessarily equally effective in foreign markets. For example, in low-income countries there is greater appreciation among consumers for sales promotions, because they like getting something for free. Competitors in a foreign country may react negatively to specific sales promotions used by the firm, and may sometimes take legal action against certain methods used. In addition, the type and size of the sales promotion, in many countries, is regulated by foreign legislation (e.g., complete prohibition of competitions), while cultural factors may necessitate changes in certain sales promotions (e.g., nature of the gift provided). Finally, in some countries, retailers are not very cooperative as regards distributing and/or handling certain sales promotions (e.g., lack of redemption centers for coupons), thus making their use problematic.

Public relations is another useful promotional tool, especially when the exporter enters a foreign market for the first time and wants to create a favorable image and publicity. Public relations may also assist in cultivating good relationships with host governments, especially in countries where foreign business involvement is viewed with suspicion. The nature of public relations should accord with the specific communication needs of each country. For instance, public speeches by the firm's managers may be more suitable in countries where relationships with government officials are important, while contributing to philanthropic events is more applicable to societies where there is distrust of foreign organizations.

Personal selling may prove particularly useful for the firm in the case of selling to countries where advertising and/or other promotional tools face legislative restrictions. However, personal selling is particularly culturally bound because it involves close personal contact and, as such, selling techniques and approaches have to be adapted to the character of each foreign market. Accordingly, the recruitment, training, motivation, and control of the sales force should fit the requirements of each country. Personal selling is particularly important in the case of exporting industrial goods, especially those that are highly priced, where explanations concerning technical issues and interpersonal relationships are vital to closing the sale (*see* GLOBAL SALES MANAGEMENT).

Finally, certain types of *direct marketing*, such as direct mail, telemarketing, and television marketing, may not be suitable for exporting to some countries, owing to poor communications infrastructure, high illiteracy rates, or incomplete/inexistent databases. On the other hand, the growing penetration of the Internet in many countries makes the use of electronic methods for contacting overseas representatives and end users an attractive option for the exporter.

Export Organization and Control

The effective and efficient execution of the export marketing strategy strongly depends on the existence of an adequate organization, and the way export activity is controlled, in order to keep strategy on the right lines (Branch, 1984; Nooman, 1985).

Export organization. To be able to coordinate and direct its foreign operations, the firm has to maintain a separate export department, staffed with adequate personnel, and managed by capable people. The presence of such a

department will give rise to continuous interest, motivation, and commitment within the firm as regards properly implementing the export marketing strategy. It will also help to carry out the specialized administrative tasks required to perform the exporting function, such as export documentation, handling of shipments, and communicating with overseas customers. Moreover, it will help to make better use of the exporting skills available within the firm, as well as facilitate training on export-related issues. If the firm opts to operate in many countries, then it is important to divide the world market into regions and appoint managers for each region.

Table 1 Strategic export marketing plan overview.

Export situation analysis
–*Internal company analysis* – Identify the firm's strengths and weaknesses in terms of export-related resources, capabilities, and other aspects
–*Export task environment* – Analyze foreign market characteristics, buyer behavior, competition, and marketing intermediaries
–*Foreign macroenvironment* – Investigate the physical, demographic, economic, sociocultural, political–legal, and technological environment of foreign countries
Export target marketing
–*Foreign country selection* – Select the most appropriate foreign countries, decide the number of countries to serve, and prioritize the countries to enter
–*Country target groups* – Segment each foreign market selected, identify the most attractive segments, and define positioning theme for each market segment
Export objectives
–*Country objectives* – Set overall objectives for each foreign country, such as market position, sales, and profits
–*Segment objectives* – Set specific objectives for each segment within each foreign market, such as market share, profitability, and customer satisfaction
Export marketing mix
–*Product* – Define export product design/style, quality, packaging, labeling, warranties, and after-sales service
–*Price* – Set export prices and end-user prices, sales terms, credit terms, and quotation currency
–*Distribution* – Establish foreign market representation, design channels of distribution, and set ways of managing channels
–*Logistics* – Design export order processing procedures, transportation methods, warehousing facilities, and inventory systems to serve foreign markets
–*Promotion* – Indicate actions for advertising, sales promotion, public relations, personal selling, and direct marketing in export markets
Export organization and control
–*Export organization* – Set up and staff the export department and assign geographic regions to specific managers
–*Export control* – Establish export-related standards, examine deviations from standards, and take corrective measures.
Contingency plans
–Indicate planned actions if events in foreign markets differ from those assumed in the exporting plan

Such geographic organization is useful in better introducing new products and services, building up distribution facilities, and training foreign sales representatives.

Export control. The firm should set up a mechanism for controlling the implementation of its export marketing strategy. First, clearly defined, easily understood, and commonly agreeable standards have to be set. These should relate to the stated exporting objectives, and be broken down for each unit and individual involved in the exporting activity. Exporting standards should take into account the specific conditions prevailing in the foreign countries targeted, such as economic situation, foreign exchange rates, and market characteristics. To establish the extent to which the various standards are met, it is important to obtain timely, reliable, and accurate information from various sources, such as internal records, strategic intelligence systems, and special survey reports. Finally, the exporter should be willing to take corrective measures, through, for example, a redefinition of the exporting objectives, a retraining or reassignment of personal roles, and the provision of more input, clarification, and guidance regarding various strategic aspects of exporting.

Concluding Remarks

To achieve success in international markets, the exporting firm has to carefully and systematically go through all steps of the strategic planning process identified earlier (for an overview of the major issues involved in each step see Table 1). As opposed to marketing strategies developed for the domestic market, export marketing strategies encounter more difficulties in design, implementation, and control. These can be attributed to (i) the stronger pressures exerted on the firm, caused by the demand of excessive company resources and specialized capabilities pertaining to exporting; (ii) the greater complexity involved, associated with the multiplicity, volatility, and heterogeneity of the international marketing environment; and (iii) the more acute business risks faced, resulting from geographic distance, fierce competition, and cultural differences in overseas markets.

Irrespective of these adversities, exporting firms should clearly understand that an export market-driven strategy is of paramount importance to achieving high performance and long-term presence in a fast-changing, globalized world. Hence, the exporter should carefully monitor foreign market trends and buyer needs, establish an organizational culture geared toward offering superior value to customers, and engage specific resources and capabilities in satisfying customers better than the competition. All these are essential in reaping the numerous financial and nonfinancial benefits accruing from exporting, which, on the one hand, will help the firm to strengthen its competitive position in the domestic marketplace and, on the other, will encourage its gradual progression to more advanced stages of internationalization.

Bibliography

Albaum, G., and Duerr, E. (2008) *International Marketing and Export Management*, 6th edn, Prentice Hall, London.

Branch, A.E. (1984) *Elements of Export Marketing and Management*, Chapman & Hall Ltd, London.

Cavusgil, S.T. and Zou, S. (1994) Marketing strategy–performance relationship: an investigation of the empirical link in export market ventures. *Journal of Marketing*, **58** (1), 1–21.

Cavusgil, S.T., Zou, S., and Naidu, G.M. (1993) Product and promotion adaptation in export ventures: an empirical investigation. *Journal of International Business Studies*, **24** (3), 479–506.

Czinkota, M.R. and Ronkainen, I.R. (2006) *International Marketing*, Harcourt, Inc., Belmont.

Hibbert, E.P. (1985) *The Principles and Practice of Export Marketing*, William Heinemann Ltd, London.

Katsikeas, C.S. and Leonidou, L.C. (1996) Export marketing expansion strategy: differences between market concentration and market spreading. *Journal of Marketing Management*, **12** (1-3), 113–134.

Katsikeas, C.S., Morgan, N., and Kaleka, A. (2004) Drivers of export venture performance: a theoretical model and empirical assessment. *Journal of Marketing*, **67** (1), 90–108.

Kotabe, M. and Helsen, K. (2001) *Global Marketing Management*, John Wiley & Sons, Inc., New York.

Leonidou, L.C. (1998) Organizational determinants of exporting: conceptual, methodological, and empirical insights. *Management International Review*, **38** (1), 7–52.

Leonidou, L.C., Kaminarides, J.S., and Panayides, P. (2007) The international marketing environment: textbook content versus academics' views. *Journal of Teaching in International Business*, **18** (2-3), 101–131.

Leonidou, L.C., Katsikeas, C.S., and Piercy, N.F. (1998) Identifying managerial influences on exporting: past research and future directions. *Journal of International Marketing*, **6** (2), 74–102.

Leonidou, L.C., Katsikeas, C.S., and Samiee, S. (2002) Marketing strategy determinants of export performance: a meta-analysis. *Journal of Business Research*, **55** (1), 51–67.

Majaro, S. (1982) *International Marketing: A Strategic Approach to World Markets*, George Allen & Unwin, London.

Nooman, C. (1985) *Practical Export Management*, George Allen & Unwin, London.

international pricing objectives and strategies

Matthew B. Myers and Marcel M. Zondag

Introduction

The price paid for a good or service represents the value the buyer places on the offering and as such price presents the one opportunity for a firm to capture customer value in tangible form and still meet its underlying goal of wealth generation. Price is recognized as a strategic tool as well as a dimension of tactical, day-to-day business decision making (Myers, 1997). In business-to-business dyadic relationships, price also acts as a relationship tool, an investment instrument of sorts, establishing trust and commitment between supply chain partners (Voeth and Herbst, 2006). On a different dimension, price functions as the vehicle of economic exchange, an integral part of the foundation and working of supply chains and, in turn, the entirety of the market economy system. Because of this multidimensional character, pricing decisions are a crucial factor in relationship marketing, that is, establishing and maintaining long-term, mutually beneficial business relationships with customers and with suppliers. The positive effects on firm performance of customer satisfaction and retention are widely recognized, so firms focus on creating a portfolio of situation and context specific relationships, allowing them to offer products and services that the customers will perceive as having a higher value than those offered by the competition (Golicic and Mentzer, 2006; Min, Mentzer, and Ladd, 2007). Price is therefore a crucial part of the integrated marketing mix as an expression of all dimensions of value: customers' perception, opportunity or switching costs, affective factors, and so on. Price provides immediate market feedback if the efforts for differentiation contained within all elements of the marketing mix have succeeded, and is the strategic lever in the marketing mix that, in return for a relative low investment, generates the highest, directly measurable returns. From a sample of Fortune 500 companies, Hinterhuber (2004) shows that a 5% price increases on average leads to a 22% increase in earnings before interest and taxes (EBIT), a much higher impact than similar percentage increase in sales volume or reduction in cost would ever have. Price is thus the one element of the marketing mix that functions as both a marketing and a measurement tool.

In spite of its obvious importance, pricing is largely neglected in the marketing and global business literature. Specifically on the issue of international business-to-business sales, the subject of this article,[1] only a limited number of works can be found (see Table 1). In this article, we will first discuss the internal and external influence factors on a firm's pricing strategies overseas, followed by a discussion on pricing as a strategic value-management tool, before we conclude with some final thoughts on the lack of theory development and empirically sound research concerning international business-to-business pricing. We will see that because of the common emphasis on consumer and behavioral pricing theories, there remains a considerable "gap" in the body of knowledge.

The Present Condition of Global Pricing Strategies

Pricing, as part of the integrated marketing mix, presents a strategic level decision for managers. In turn, price is the opportunity for a firm to capture the hard-to-measure concept of customer value. Customer value is often expressed in terms of economic value, a construct that adds the price of the customer's best alternative product or service (the reference value) and the additional positive or negative value the customer places on those attributes of the firms product or service that differentiate it from the so referenced product or service (Nagle and Holden, 1995, p.74). The economic value analysis presents the firm with the possible price level it can achieve. However, this is only part of the pricing strategy; cost, profit goals, and competitor pricing are other factors to be weighed in the process of determining the obtainable price range for a specific product or service (Hinterhuber, 2004). The wide variety of reference values found in overseas markets makes this an extremely important yet complex activity for global managers. Yet, the

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 International business-to-business pricing studies.

Cite	*Study/Topic*	*Results/Conclusions*
(Myers, 1997)	Empirical study of US-based export managers Topic: Manager attitude toward pricing policies.	Owing to cost constraints and lack of communication and coordination between functional departments, international pricing policies are often ad hoc and cost-based only.
(Tzokas *et al*. 2000)	Empirical study of UK-based exporting manufacturers Topic: internal antecedents of firm orientation toward strategic export pricing. *Firm size is not considered as a variable in this study*.	Firms with a long-term customer focus and a positive attitude toward exporting tend to utilize "Strategic Export Pricing;" more formalized and centralized price management to maximize export returns.
(Myers and Harvey, 2001)	Empirical study of US-based exporting manufacturers Topic: internal and external antecedents of export pricing-control decisions. *Study conceptualizes a dynamic model*.	Larger, more experienced firm will maintain export pricing control over intermediaries, this is however not conducive to long-term export relationships; export pricing decisions should be made by entities closer to the customers. Export pricing decisions affect the strategic performance of the firm.
(Cavusgil, Chan and Zhang, 2003)	Empirical study of US-based exporting manufacturers. Topic: Firms' Strategic Orientation toward Export Pricing ('SOEP')	Classifying by SOEP, 4 firm typologies are recognized: (i) highly centralized experienced exporters, (ii) decentralized experienced exporters, (iii) new to export, and (4) experienced poor performers. Only the first two types of firms are successful as they adhere to a pricing strategy that matches their particular export environment.
(Bolton and Myers, 2003)	Empirical study of IT service contract customers in Asia, Europe, and North America. Topic: Price elasticity measured across market segments within national borders and across borders, influence of national/cultural preferences	Price can be used as a market segmentation tool; both vertical (country specific) and across borders on a regional *and* global scale as business customers are shown to have similar price elasticity across markets.

(*continued overleaf*)

Table 1 (*continued.*)

Cite	*Study/Topic*	*Results/Conclusions*
(Myers, 2004)	Empirical study of US-based exporting manufacturers. Topic: The effect of congruence or "fit" between export pricing strategy and venture strategy on firm performance	Not matching export pricing strategies with overall firm objectives, a common occurrence, will lead to suboptimal firm performance.
(Lancioni, 2005)	Propositional article Topic: Normative discussion about indus Empirical study of Portugal-based multi-industry exporting firms trial pricing.	In order to align pricing strategy with firm objectives a detailed procedure to formulate a pricing plan as integral part of the firm marketing plan is presented.
(Sousa and Bradley, 2008)	Topic: Internal and external antecedents of price adaptation and export performance.	Price adaption in reaction to environmental factors is common. Export experience and the number of markets serviced will lead to more rigid (centralized) pricing. Export performance is better in market environments which is much different from the home market.

pricing decision, or pricing strategy, should be at the core of any industrial marketing plan (Myers, 2004; Lancioni, 2005). In many instances however, export pricing plans are mere derivatives of domestic pricing, sidestepping the intricacies of conducting international business and leaving firms with similar pricing and pricing structures across foreign markets (Myers, 1997). Because of the idiosyncratic differences between each export market, and even between individual industrial exchange relationships, many exporting firms end up underpricing, leaving "money on the table," or overpricing, losing market share to local competitors. The latter also carries with it the risk of facilitating the creation of gray markets, the unauthorized transfer of goods from low-, to higher price markets, cannibalizing the firm's revenues.

Before addressing these anomalies of the export pricing process, firms will first need to better understand the external and internal factors that affect their export pricing decisions. These are outlined below.

Factors external to the organization.

Understanding the global customer. It is a truism that a firm needs to understand who its customers are or, more so in the context of this discussion, the economic value customers attach to the firm's offerings. Determining economic value involves measuring and interpreting the price effects of such diverse factors as alternative goods or services, budget restraints, switching costs, and a wide array of affective and cognitive evaluations that customers make in making their buying decision. Understanding the value perception of the customer will allow to approximate price sensitivity of different customer groupings, allowing for market and price segmentation (Bolton and Myers, 2003), or providing the benchmark for measuring the effectiveness of other marketing mix elements such as product (development) and promotion. It is similarly true that the geographic, regulatory, and cultural "distance" affecting cross-border transactions will add complexity to this information gathering process, leaving export managers with an incomplete set of

parameters on which to base their pricing strategies (Samli and Jacobs, 1993).

Although the industrial buyer fulfills a role not dissimilar to the individual consumer (Alderson, 1965), the actual industrial buying process differs greatly from consumer buying behavior; the decision is often made by groups of people (or decision making unit) to satisfy the needs and objectives of different constituents within the firm, the process if often more formalized than in consumer buying, and the personal and organizational risks associated with the purchase decision are often much greater than in the case of consumer buying (Moriarty, 1983). The majority of pricing research, however, is concerned with consumer buyer behavior and thus has little relevance and applicability in the field of international business-to-business pricing, and to a large extent our understanding of effective pricing to global customers suffers from a paucity of research in this area.

Heightened global competition. Establishing economic value is already an exercise in competitive analysis, comparing one's product and services against acceptable alternatives. The operating term here is "acceptable," signaling the need to include products or services in the analysis that may fall well outside of the often myopic view managers have about the positive attributes of their own products or services. As a consequence of this myopia, firms often find themselves facing unexpected competition, specifically in international markets where familiarity with competitors is limited relative to the knowledge of the competition in the home market.

Acquiring information about competitive products and pricing is mostly left to the sales force or, in the case of international business, to local distributors or sales agents and is therefore much less based on independent external sources than one would assume (Stottinger, 2001). There are two inherent problems with this practice. First, the sales function is mostly focused and rewarded on a firm's top-line sales and looks to set pricing as low as possible to book more sales, commonly mistaking "customer value" with "customer satisfaction," given that the latter will eventually increase if prices are low enough. Secondly, pricing intelligence is often directly provided by the customers' purchasing staff, a business function that itself is looking to minimize prices, casting doubt on the validity of the data.

Diversity of overseas markets. Selling outside of the relative familiarity of the home market is wrought with logistical, regulatory, and cultural influences that most firms are ill prepared for. Certain jurisdictions may for instance levy high import duties, or restrict the free flow of goods for sociopolitical or protectionist reasons. Technical requirements and labeling laws may differ widely, increasing a firm's cost structure. Most firms try to minimize complexity by standardizing products, services, and business processes and will try to implement a uniform pricing policy across markets. As a result, firms select export markets and customers resembling their domestic market and customer, and do so in a more opportunistic manner than one would expect (Bilkey and Tesar, 1977; Kogut and Singh, 1988; Maignan and Lukas, 1997). As export experience grows, so does the firm's ability to make more deliberate pricing decisions. Before that, most price decisions are ad hoc adaptations of existing pricing, anecdotally driven. In order to circumvent the difficulties of exporting, smaller and/or lesser experienced firms often rely on export intermediaries such as stocking distributors in their exporting efforts. For these companies export activity essentially ends with delivery to the intermediary. With the exception of sales and promotional support, these exporters often leave pricing control to the intermediary (Myers and Harvey, 2001). This tendency to suboptimize export performance and revert to traditional pricing models is the result of a predominant internal focus of many firms when it comes to pricing decisions, discussed in the next section.

Factors internal to the organization. The importance of external information in the pricing process cannot be emphasized enough. The concept of customer value or "putting the customer first" is based on understanding the totality of customers' economic value, inclusive of all the competitive and environmental factors. Indeed, most firms are externally focused in

deciding on their integrated marketing strategy, but empirical research shows that firms still tend to focus internally when making pricing decisions, paying only lip service to value pricing (Nagle, 1987; Myers, 1997). Manufacturing cost, international sales costs, and profit margins as defined under accounting rules are the factors most influencing export pricing decisions (Solberg, Stottinger, and Yaprak, 2006). Most export firms use a cost-plus pricing model, adding desired gross margin to unit cost as determined by the accounting function and based on the firm's past expenditures. More experienced exporters do often pay more attention to the "voice of the customer" and operate a flexible cost-plus pricing model, making ad hoc pricing adaptations to counter competitive pricing, gain market share, or account for late payment, taxes, or fees.

Extant literature presents numerous other pricing strategies and tactics, often derived from traditional cost-plus models. The fundamental problem with using traditional cost accounting (or management accounting) for making pricing decisions is that it uses historical data to create a simplified representation of an inherently dynamic concept (cost per unit). For instance, price directly influences sales volume, which in turn should immediately affect the allocation of fixed costs to each unit downward (Nagle, 1987). More recent management accounting practices such as activity-based costing are still inwardly focused on cost instead on customer value and as such offer little improvement (Baker, 2006). Beside cost accounting practices, financial reporting standards do not allow for dynamic adaptation, and also entice managers with a seemingly sequential tabulation of pricing elements (costs and margin) that does not account for any relevant external market intelligence, and reduces the determination of customer value to an internal discussion (most often between the accounting and sales functions) about what constitutes a "fair" margin. The trends toward global standardization of financial reporting through the efforts of the International Accounting Standards Board (IASB) is not likely to change this, as historical cost remains as one of the prevalent basis for financial reporting. We argue that this fundamental problem can only be addressed by a structural change in the formulation and execution of pricing strategy, requiring a shift in business philosophy toward value creation that may well extend to changes in organizational structure.

Pricing as Value Management

Baker (2006), Nagle and Holden (1995), and Hinterhuber (2004) present persuasive arguments that internally focused or "cost-plus" pricing models often lead to suboptimal performance as firms seek to increase throughput at full capacity, thus reducing the cost per unit. However, recent events indicate that market success is not based on market share and size (see General Motors' loss of US market position) but on a firm's ability to generate continuing profitable sales by offering superior economic value. Achieving the highest level of "lean," efficient production throughput will have minimum benefits if throughput is of no economic value to the customers. Instead, the focus should be on the marketplace, looking to establish what products and services the firm can effectively produce that match the wants and needs of its customers as well as determining what the perceived economic value of these offerings are from the viewpoint of the customer.

The "power" of such customer focus is underlined by recent research (Bolton and Myers, 2003), showing how customers' price elasticity of demand transcends national borders, holding that pricing strategies can not only segment national markets (vertical segmentation), but also delineate market segment across national borders, (horizontal segmentation) – further evidence of the paramount role for pricing in the integrated marketing strategy.

All of this points to a relatively limited extant literature on the subject of business-to-business export pricing, creating a considerable "gap" in the body of knowledge.

"Mind the Gap": Issues in Export Pricing Practice and Theory

Despite repeated calls for more empirical research and theory building, export pricing, specifically industrial export pricing, remains

an understudied subject in marketing. Since *Industrial Marketing Management's* 2005 special issue on industrial pricing only a handful of scholarly works have been published. Intraorganizational transfer pricing, pricing controls, or consumer oriented single market issues seem to be more popular research subjects than intercompany export pricing. Following Solberg, Stottinger and Yaprak (2006), we agree that the many complicating organizational and environmental factors could well lead to this shortfall. Given today's hyper-competitive environment this is rather surprising, taking into account the immediate impact of price on export performance and profitability, and in the light of the prevalence of price-related globalization issues discussed in the popular business press.

In many cases exporting is more opportunistic than one would think or expect (Andersen and Buvik, 2002), and less the result of a well-executed strategic marketing effort. Because of the initial opportunistic character of export transactions, firms tend to have heightened expectations about the profitability of additional export efforts without sufficient appreciation for the complex, organization-wide impact of formulating and executing a truly integrated marketing strategy for creating and capturing cross-border customer value (Flint, 2004). In addition, in many cases firms reduce export sales costs by engaging export intermediaries, never establishing the dedication and business philosophy to manage export activities for maximum customer value and firm financial performance. Contrary to espoused commitment to long-term partnership with customers and suppliers, it is common for employee remuneration systems to be based on short-term financial goals, offering managers little motivation to pursue integrated long-term exporting strategies instead of taking a shorter term transactional approach. As a result, pricing in overseas markets to take advantage of customer values (as opposed to short-term profits) is rare. This often culminates in dysfunctional export business relationships. Irrespective of their purported financial success, these relationships are often perceived as stressful and risky, further diminishing the motivation for firms to achieve their full export potential. Future research on the detriments of short-term pricing strategies on long-term export partnerships is needed to address the problems associated with extrapolating cost-based pricing models to overseas markets.

We argue that firms will need to institutionalize two key components to their global pricing efforts. First, firms often assume that the existence of an export (sales) function in an organization automatically denotes the ability to effectively gather and interpret market data (Myers, 1997; Hinterhuber, 2004), and/or the ability to deal with opportunistic behavior of one's foreign distributor (Cavusgil, Deligonul, and Zhang, 2004). Unfortunately, this is not possible without a concerted effort to gather market-related data which enables the effective setting of prices. The importance of gathering and interpreting (sense-making) such customer, market, and partner intelligence is further underlined by recent research, again confirming that an increase in economic value allows for proportionate price increase (Hulten, Vistrom, and Mejtoft, 2009), and often firms with export experience and suitably flexible pricing strategies achieve better export performance (Sousa and Bradley, 2008).

Second, firms must develop a strategic export pricing plan congruent with the overall firm-level strategic plan, that captures all necessary customer, competitor, and market intelligence and sets forth how the firm's pricing plans will contribute to attaining the firm's goals and objectives (Myers, 2004; Lancioni, 2005). In many cases executing such a integrative and dynamic pricing strategy will require organizational changes to facilitate sound pricing decision making, instead of ad hoc and intuitive adaptations of cost-plus pricing as a result of intraorganizational conflict between accounting and sales functions (Hinterhuber, 2004), for instance through the creation and empowerment of a corporate pricing function. While understanding the difficulties involved in these types of organizational changes, these shifts are often the only manner in which pricing plays a central role in export management and serves as a source for sustainable competitive advantage (Stottinger, 2001; Theodosiou and Katsikeas, 2001).

Against the backdrop of global business realities, it is clear why there have been repeated calls for timely and relevant export pricing research. Indeed, most international pricing research is descriptive, concentrating on specific price factors, industries, or geographical markets, resulting in findings of limited external validity. Findings are often cross-sectional, self-reporting surveys, raising questions about the causality captured in the conceptual models presented. The linkages between pricing strategies and export performance is often measured differently across studies raising questions regarding the validity of some of the research. Furthermore, integrative conceptual models (addressing multiple markets, industries, outcomes, and pricing strategies) are yet to be fully developed. One could argue that export pricing theory is (slowly) being developed "bottom up" as opposed to "top down." On a more positive note, we consider it a valuable development that information gathering and knowledge sharing are becoming more prevalent in pricing studies. Also, we welcome the inclusion of the strategic, organizational, and human interaction aspects of export transactions in more recent empirical research and theory building literature. We do however feel that many of the rigor and relevance issues surrounding industrial export pricing research are a direct consequence of the considerable difference between espoused and actual pricing practices and a social/organizational desirability bias in the research measurements. When we more carefully review the successful global pricing strategies practiced by gold-standard marketing managers, we note that recipes for success are almost all based on creating and capturing superior customer value. In other words, as long as scholars provide a "quick fix" global pricing solution sets without accepting the need for a fundamental paradigm shift toward longer term, value-oriented pricing strategies, our understanding of the global pricing function will remain limited.

Conclusions

Pricing is the one marketing tool with the unique ability to indicate, dictate, and capture customer value. As such it plays a critical role in the formulation and execution of export strategy. Firms (and scholars) should realize the "execution premium" (Kaplan and Norton, 2008) that is available from moving beyond traditional cost-based pricing methods and instead implementing a comprehensive information and knowledge acquisition strategy, leading to the formulation of a strategic pricing plan based on customer perceived economic value. Similarly, continuing consolidation, globalization, and transparency of markets will amplify the negative effects of not having a comprehensive, long term, value-based pricing strategy on the financial performance of exporting firms.

Endnotes

[1] We have limited our discussion to export pricing, more specifically, industrial (or B2B) pricing in cross-border transactions.

Bibliography

Alderson, W. (1965) *Dynamic Marketing Behavior: A Functionalist Theory of Marketing*}, Richard D. Irwin Inc., Homewood.

Andersen, O. and Buvik, A. (2002) Firms' internationalization and alternative approaches to the international customer/market selection. *International Business Review*, **11**, 347–363.

Baker, R.J. (2006) *Pricing on Purpose: Creating and Capturing Value*, John Wiley & Sons, Inc., Hoboken.

Bilkey, W.J. and Tesar, G. (1977) The export behavior of smaller-sized wisconsin manufacturing firms. *Journal of International Business Studies*, **8** (1), 93–98.

Bolton, R.N. and Myers, M.B. (2003) Price-based global market segmentation for services. *Journal of Marketing*, **67**, 108–128.

Cavusgil, S.T. and Chan, K. and Zhang, C. (2003) Strategic orientations in export pricing: a clustering approach to create firm taxonomies. *Journal of International Marketing*, **11** (1), 47–72.

Cavusgil, S.T. and Deligonul, S. and Zhang, C. (2004) Curbing foreign distributor opportunism: an examination of trust, contracts, and the legal environment in international channel relationships. *Journal of International Marketing*, **12** (2), 7–27.

Flint, D.J. (2004) Strategic marketing in global supply chains: four challenges. *Industrial Marketing Management*, **33**, 45–50.

Golicic, S.L. and Mentzer, J.T. (2006) An empirical examination of relationship magnitude. *Journal of Business Logistics*, **27** (1), 81–108.

Hinterhuber, A. (2004) Towards value-based pricing – an integrative framework for decision making. *Industrial Marketing Management*, **33**, 765–778.

Hulten, P., Vistrom, M. and Mejtoft, T. (2009) New printing technology and pricing. *Industrial Marketing Management*, **38** (3), 253–262.

Kaplan, R.S. and Norton, D.P. (2008) *The Execution Premium: Linking Strategy to Operations for Competitive Advantage*, Harvard Business Press, Boston.

Kogut, B. and Singh, H. (1988) The effects of national culture on the choice of entry mode. *Journal of International Business Studies*, **19** (3), 411–432.

Lancioni, R.A. (2005) A strategic approach to industrial product pricing: the pricing plan. *Industrial Marketing Management*, **34**, 177–183.

Maignan, I. and Lukas, B.A. (1997) Entry mode decisions: the role of managers' mental model. *Journal of Global Marketing*, **10** (4), 7–22.

Min, S. and Mentzer, J.T. and Ladd, R.T. (2007) A market orientation in supply chain management. *Journal of the Academy of Marketing Sciences*, **35**, 507.

Moriarty, R.T. (1983) *Industrial Buying Behavior*, D.C. Heath and Company, Lexington

Myers, M.B. (1997) The pricing of export products: why aren't managers satisfied with the results. *Journal of World Business*, **32** (3), 277–289.

Myers, M.B. (2004) Implications of pricing strategy – venture strategy congruence: an application using optimal models in international context. *Journal of Business Research*, **57**, 591–600.

Myers, M.B. and Harvey, M. (2001) The value of pricing control in export channels: a governance perspective. *Journal of International Marketing*, **9** (4), 1–29.

Nagle, T.E. (1987) *The Strategy and Tactics of Pricing*, Prentice Hall, Englewood Cliffs.

Nagle, T.E. and Holden, R.K. (1995) *The Strategy and Tactics of Pricing: A Guide to Profitable Decision Making*, Prentice-Hall, Inc., Englewoods Cliffs.

Samli, A.C. and Jacobs, L.W. (1993) International pricing decisions: a diagnostic approach. *Journal of Marketing Theory and Practice*, Fall/Winter, **1** (4), 29–41.

Solberg, C.A., Stottinger, B. and Yaprak, A. (2006) A taxonomy of the pricing practices of exporting firms: evidence from Austria, Norway, and the United States. *Journal of International Marketing*, **14** (1), 23–48.

Sousa, C.M.P. and Bradley, F. (2008) Antecedents of international pricing adaptation and export performance. *Journal of World Business*, **43**, 307–320.

Stottinger, B. (2001) Strategic export pricing: a long and winding road. *Journal of International Marketing*, **9** (1), 40–63.

Theodosiou, M. and Katsikeas, C.S. (2001) Factors influencing the degree of international pricing strategy standardizations of multinational corporations. *Journal of International Marketing*, **9** (3), 1–18.

Tzokas, N. and Hart, S., Argouslidis, P.C. and Saren, M. (2000) Strategic pricing in export markets: empirical evidence from the UK. *International Business Review*, **9**, 95–117.

Voeth, M. and Herbst, U. (2006) Supply-chain pricing – a new perspective on pricing in industrial markets. *Industrial Marketing Management*, **35**, 83–90.

international marketing channels

Daniel C. Bello

Introduction

International marketing channels consist of interdependent organizations participating in business activities necessary to make a product or service available for use in markets around the world (Coughlan *et al.*, 2006). The structure or design of global channels can vary widely, comprising various organizational participants (foreign distributors, agents, subsidiaries, etc.) engaged in a diverse set of functional activities (selling to foreign customers, warehousing inventory, etc.) (Bello and Briggs, 2009). Channel designs range from completely integrated corporate channels (wholly owned subsidiaries) to closely managed partner networks (alliances; exclusive, highly cooperative middlemen) to loosely organized distribution systems (nonexclusive, arms-length middlemen). The mix of forces driving international channel design includes factors associated with the original equipment manufacturer (OEM or brand owner), channel members, and end users. Each of these key players – manufacturer, middleman, end user – places a unique set of requirements and constraints on the channel activity set – the business tasks and activities that must be conducted to provide value to end users.

Activity sets can be conducted within various channel designs, depending on the feasibility of governing-needed activities through institutional arrangements (IAs) such as contracting and ownership (Anderson and Gatignon, 1986; Zhao, Luo, and Suh, 2004). For example, a manufacturer's international channel may require that specialized warehousing, selling, and other unique activities be conducted in foreign markets (*see* DESIGNING A GLOBAL SUPPLY CHAIN: OPPORTUNITIES AND CHALLENGES; GLOBAL SALES MANAGEMENT). What is the best channel design for these necessary channel activities? Should the manufacturer contract with a local public warehouse firm and sales agency, or should it establish and own a local subsidiary, using employees to perform these tasks? Issues of channel design reflect the relative ability of IAs to govern channel activity sets in an efficient and effective manner (Bello and Zhu, 2006). Importantly, the feasibility of contracting, ownership, and other arrangements depend on the characteristics of the proposed activities and the institutional environments (IEs) of the home and host countries (Grewal and Dharwadkar, 2002). As will be shown, a complex interplay among firms, exchange characteristics, and IEs yields a set of feasible business activities and a workable channel design.

Determinants of Channel Activity Sets

Channel activity set refers to the marketing, logistic, and other business tasks and activities performed by channel participants during the process of making a product or service available for use. The value-added activities of a channel can be represented as marketing flows since work processes such as physical possession, ownership, and promotion flow forward through the channel, while others such as ordering and payment move up the channel from the end user (Coughlan *et al.*, 2006). Each basic work process involves many interdependent and costly tasks and activities. For example, in the flow of physical possession from factory to user, a host of warehousing, inventory, transport, and other logistic tasks may be necessary, each of which burdens channels members with direct costs, overhead, and business risk (*see* DESIGNING A GLOBAL SUPPLY CHAIN: OPPORTUNITIES AND CHALLENGES). As Figure 1 illustrates, the many factors determining the particular set of tasks and activities that are conducted in international channels reflect manufacturer, middleman, and user needs.

A manufacturer's marketing positioning strategy, resource endowments, degree of internationalization (*see* INTERNATIONAL ADVERTISING – IS THERE STILL A STANDARDIZATION VERSUS LOCAL ADAPTATION DEBATE?), among other factors, influence the set of activities necessary to move a firm's product and its title to end users. Marketing strategies based on product attributes run the gamut, extending from low-cost, utilitarian to high-quality, prestigious market positions; but, whatever a manufacturer chooses "the investments and activities required to realize the position must be specified" (Ghosh and John

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

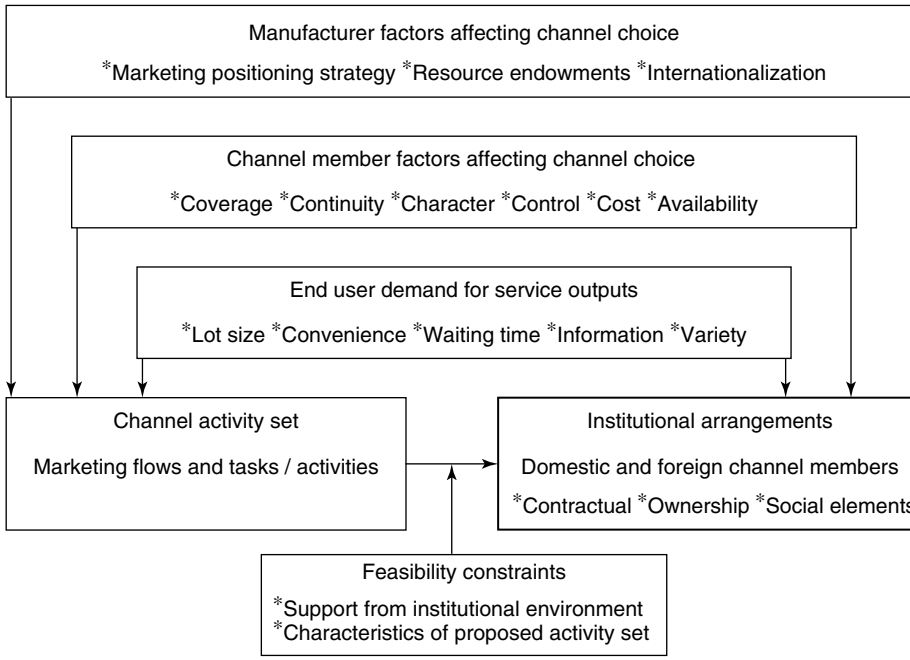


Figure 1 Determinants of international channel design.

1999, p. 143). For example, the specific activities associated with the promotional flow (*see* INTERNATIONAL ADVERTISING – IS THERE STILL A STANDARDIZATION VERSUS LOCAL ADAPTATION DEBATE?; MARKET ENTRY AND EXPANSION) toward the end user differ substantially for prestigious versus utilitarian positions: the basic nature and type of personal selling, advertising, and sales promotion tasks required at the wholesale and retail levels are unique to each positioning. Likewise, the particular resource endowment of a manufacturer influences the activity set required to support the distribution of its products (Bello and Briggs, 2009). Firms with key patents, high brand equity, extensive partner networks, and other valuable resources and capabilities will require negotiation, promotion, and other business activities from channel members that differ substantially compared to the channel requirements of under-resourced manufacturers. As for a firm's degree of internationalization, a narrow domestic or regional focus relative to a broad global focus influences a firm's overall approach to international marketing, including the particular tasks and activities required of domestic and international trading partners (Czinkota and Ronkainen, 2007).

The channel members available to participate in a manufacturer's distribution system also influence the particular mix of activities that can and will be conducted (*see* INTERNATIONAL TRADE INTERMEDIARIES). For example, a multinational corporation (MNC) may target rural, lower-income Felipino consumers, though it may lack the necessary capabilities to self-perform promotion, logistic, and payment collection tasks for this difficult-to-serve user segment. Given its resource shortcomings, the manufacturer may prefer a competent distributor as its local channel partner; yet, local distributors with the necessary sales coverage, cost structure, and complementary product lines may simply not exist. These types of channel-member factors have been summarized as the C's of international middlemen (Czinkota and Ronkainen, 2007): their coverage, continuity, character, control, cost, and so on. Thus, manufacturer factors may call for a foreign distributor, but local middleman factors may discourage this

choice, leading the firm, perhaps, to use a local sales agent supplemented by a corporate market development manager for the country.

End-user demands for channel services are often the most important drivers of required channel activities, since the ultimate success of any distribution system hinges on its ability to deliver value to the final customer (Coughlan *et al.*, 2006). As benefits for the end user, channel service outputs satisfy how users prefer to shop and reflect their purchase preferences and patterns. Essentially, user service demands answer the question "Besides the product itself, what is it that end users want from the buying process?" Users tend to prefer a specific mix of channel services such as lot size, spatial convenience, delivery or waiting time, product/brand variety, product information, after-sales services, and so on. Different segments of end users may demand substantially different channel services. Compared to their urban, high-income countrymen, rural, low-income Filipinos may require a smaller lot size (single-serve packaging), retail within walking distance, no waiting time (due to an absence of household inventories), but less brand variety and after-sale services, and so on. To target both segments of Filipino consumers, an MNC must develop different channels, each capable of performing business tasks necessary to satisfy the mix of services demanded by each customer type.

Institutional Arrangements

Regardless of factors driving a channel's activity set, the resulting mix of tasks and activities to be performed can be organized and managed within a variety of alternative channel designs (Anderson and Gatignon, 1986; Zhao, Luo, and Suh, 2004). Since channel members differ in cost efficiency and task effectiveness, they tend to specialize in performing activities yielding channel designs where a member participates in one or more flows but not in others (*see* MARKET ENTRY AND EXPANSION; STRATEGIC EXPORT MARKETING–ACHIEVING SUCCESS IN A HARSH ENVIRONMENT). In the case of a direct channel, a manufacturer with a subsidiary may participate in physical possession (operates foreign warehouses), ownership (retains title), and promotion (manages local sales force) in the foreign market. In the case of indirect channels, a manufacturer, when using foreign sales agents, may only relinquish selling while continuing to retain title and local warehouses; when using foreign distributors, a manufacturer would not participate in these flows since it sells the product to a local merchant middleman who holds title and performs local logistic and sales activities.

The allocation of tasks across firms in a channel reflects the IAs employed to ensure that participating firms conform to channel requirements by binding parties to their role responsibilities, guaranteeing economic returns for channel participation, and motivating cooperation (Bello, Lohtia, and Sangtani, 2004). IAs are the contractual, ownership, and social elements necessary to support the performance of the channel activity set within a channel design (Carson *et al.*, 1999). By establishing the rules of exchange, IAs encourage coordination and determine a division of gains sufficient to motivate all firms to participate in their assigned channel flows. In ongoing distribution systems, it is the contracts, ownership, and social forces that induce sufficient cooperation among self-interested firms to ensure each independently implements the investments and activities required to deliver value to end users (*see* INTERNATIONAL RELATIONSHIP MARKETING; MARKETING STRATEGY IMPLEMENTATION).

Feasibility constraints. As noted in Figure 1, the feasibility of managing channel activities through contracts, ownership, and social elements depends on the characteristics of both the proposed activity set and the IE in the home and foreign market (Carson *et al.*, 1999). While value-added activities can occur within a variety of direct and indirect channel configurations, distribution activities are always conducted through exchange transactions between a principal (actor needing an activity to be conducted) and an agent (actor conducting an activity). An OEM principal requiring local selling and delivery for its product in a foreign market may use a corporate hierarchy or subsidiary (ownership element of the IA): in such direct channels, the principal engages in

an exchange transaction with a subsidiary-based employee-agent, the actor conducting the activities in the foreign market. Alternatively, the principal may use a market-based exchange with a foreign distributor (contract element of the IA): in indirect channels, the principal's exchange transaction occurs within a competitive market since the distributor-agent in this case is selected from among several alternative, independent merchant middlemen available in the foreign country (*see* INTERNATIONAL NEGOTIATIONS).

Characteristics of the activity set. Much theorizing has focused on determining which element of an IA – market contract, hierarchical ownership, or social/relational – is best suited to govern exchange transactions for global distribution activities (see Bello and Briggs, 2009). While various theories provide important insights to channel design, transaction cost analysis (TCA) is the dominant paradigm for analyzing institutional structure and the efficient organization of business activities (Rindfleisch and Heide, 1997, Williamson, 1996).

TCA employs an economic efficiency criterion for channel design: efficient IAs minimize the transaction costs associated with conducting activities by aligning the attributes of exchange with the governance abilities of alterative IA elements (Bello and Zhu, 2006; Williamson, 1996). Transaction costs are the costs of managing an exchange transaction, including ex ante costs (drafting and negotiating agreements) and ex post costs (monitoring and enforcing agreements) (Rindfleisch and Heide, 1997). In other words, "TCA logic focuses on the characteristics of exchanges involving distribution activities because, given the potential of partner opportunism, exchange attributes may give rise to the governance problems of safeguarding specific investments, adapting to change, and evaluating performance" (Bello and Briggs, 2009, p. 398). A principal–agent exchange is characterized by high transaction costs to the extent a channel activity (i) requires specific, rather than general, assets, (ii) is subject to unanticipated changes in future circumstances (environmental volatility), and (iii) involves performance ambiguity or difficulty in assessing task outcomes (internal uncertainty). Specific assets create safeguarding problems since an investor becomes locked in to the relationship: investments dedicated to a channel partner have little value in alternative uses, subjecting the investor to exploitation of profits by an opportunistic partner. Environmental volatility causes adaptation problems by creating unanticipated opportunities for a partner to exploit change by opportunistically renegotiating terms of trade. Performance ambiguity gives rise to measurement problems, making it difficult and costly for a principal (or a third party such as a court) to evaluate whether the agent's conduct is in compliance with agreements.

High transaction costs impose significant constraints on the feasible IA elements capable of efficiently governing the principal–agent exchange for an activity. For example, a manufacturer developing a foreign market may require specialized warehousing (unique material handling, special temperature/humidity control, etc.) for its product, and may prefer to address these needs by contracting with a local public warehousing firm (*see* DESIGNING A GLOBAL SUPPLY CHAIN: OPPORTUNITIES AND CHALLENGES). However, high transaction costs make this channel activity noncontractible since a local warehousing firm contemplating modifying its warehouse to service the manufacturer would be exposed to safeguarding and adaptation problems that are inadequately solved through simple contracting. First, specialized warehousing investments only have value in the service of this specific principal, making the warehousing firm vulnerable and unable to retaliate against bad principal behavior because of its lack of alternatives. Hence, profits or economic rents from the warehousing firm's investments cannot be guaranteed by a contract since (i) during the contract term the principal can exploit the agent (extra demands, slow payments, etc.) because of the agent's locked-in situation, and (ii) upon expiration the principal may opportunistically demand unfair terms to continue the relationship. Second, the market-development effort may prove volatile, failing to provide the steady volume necessary to operate the warehouse at capacity, saddling the warehousing firm with unexpectedly high costs and no profits. Given these characteristics, a simple contract does not solve the warehousing

firm's safeguarding and adaptation problems in a manner sufficient to induce participation in the manufacturer's channel system.

While high transaction costs create contracting difficulties, low transaction cost situations are easily governed through simple contracting (Williamson, 1996). A manufacturer requiring general warehousing (employing common, industry-standard assets) for a product with a proven sales record in a foreign market will encounter little difficulty in contracting with a local warehousing firm (*see* DESIGNING A GLOBAL SUPPLY CHAIN: OPPORTUNITIES AND CHALLENGES). Low transaction costs evidenced by the lack of safeguarding and adaptation problems make this channel activity easily contractible. With a simple contract, a warehousing firm can confidently participate in this manufacturer's channel system. In contrast, the manufacturer requiring specialized warehousing must motivate participation by offering increasingly complex contracts, perhaps addressing safeguarding issues through clauses guaranteeing renewal at attractive terms or addressing adaptation problems by guaranteeing minimum payments regardless of volume. At some point, however, patching together complex clauses fails to motivate participation since governance problems remain unresolved.

The contractible subset of the channel activity set is composed of channel activities subject to contractual solutions to governance problems (Carson *et al.*, 1999). For noncontractible activities, the IA element of ownership may be better aligned with the governance problems characterizing these principal–agent exchanges. The manufacturer requiring specialized warehousing may have to own the unique material-handling and humidity equipment located at the foreign facility, and be responsible for repairing and maintaining the assets it owns in the warehouse. In other words, the warehousing firm's participation is obtained by devising a suitable IA for each aspect of the physical possession flow: a five-year contract for the contractible activities (leasing warehouse space) and manufacturer ownership of noncontractible activities (specialized equipment). Such ownership rearrangement of high transaction cost activities reflects the TCA prescription of market governance (contracting) as the default IA for nonproblematic activities, while hierarchical governance (ownership) is used to efficiently address exchanges posing serious safeguarding, adaptation, and measurement problems (Rindfleisch and Heide, 1997).

In this way, an appropriately designed IA is necessary for every aspect of the channel activity set (Bello and Zhu, 2006). In terms of the promotion flow (*see* INTERNATIONAL ADVERTISING – IS THERE STILL A STANDARDIZATION VERSUS LOCAL ADAPTATION DEBATE?), for example, a manufacturer may decide to sell a complex, technical product in a foreign market. What is the appropriate IA for the foreign selling activities? If the principal–agent exchange for the local selling effort invokes few safeguarding, adaptation, and measurement concerns, a simple contract with a foreign sales agency is the efficient channel design. However, owing to its inherent complexity and frequent innovations, if the product requires specialized learning to sell effectively, a sales agency may be reluctant to invest in extensive and continuous training (a specialized human asset, nonredeployable to other principals). Fearful of lowered commissions or even termination after it develops the foreign market for the product, the sales agency may decline participation in the channel, worrying that economic return from its investments cannot be sufficiently guaranteed through a contract. The manufacturer may find it necessary to bring the selling task in-house, establishing a sales subsidiary, and training its own salespeople.

Social elements (informal rules such as relational norms) are the final IA mode, which are most effectively employed for activities subject to difficulties in both contracting and ownership (Carson *et al.*, 1999). Since relationship norms (*see* INTERNATIONAL RELATIONSHIP MARKETING) employ social enforcement to bind parties to their channel activities, difficult exchanges can be governed by certain relational norms (Bello, Lohtia, and Sangtani, 2004). As shared understandings between the parties, norms such as information sharing (willingness to exchange proprietary information), flexibility (willingness to adapt procedures), and solidarity (desire to maintain relationship) can enforce compliance to social expectations regarding

assigned activities. Returning to the specialized warehouse example, repairing and maintaining unique material handling and humidity equipment may involve tacit knowledge that is so causally ambiguous and embedded in the manufacturer's corporate routines that it cannot easily be transferred to the foreign warehousing firm. What kind of IA can best govern the costly and difficult maintenance of the manufacturer-owned equipment in the leased space of a foreign warehouse? While manufacturer technicians (ownership) can periodically visit the facility, repair and maintenance requires a great deal of information sharing, flexibility, and solidarity by the partners to make such visits timely and effective. Without strong relational norms, the contractual and ownership elements by themselves may not be adequate to sustain the specialized foreign warehousing activity for the manufacturer's international channel (*see* INTERNATIONAL RELATIONSHIP MARKETING).

Support from the institutional environment. As macrolevel aspects of society (*see* MARKETING ASPECTS OF PSYCHIC DISTANCE; MARKETING ASPECTS OF CULTURAL DISTANCE), the IE consists of underlying and distinguishing elements of a country's setting and context (Bello, Lohtia, and Sangtani, 2004). Theorists (Scott, 2001) recognize three elements composing the IE: regulative, normative, and cultural-cognitive forces. Importantly, differing IEs around the world (*see* BASE OF THE PYRAMID MARKETS: CULTURE INSIGHTS AND MARKETING IMPLICATIONS; EMERGING MARKETS) moderate the ability of exchange partners to develop the IAs necessary to support a desired channel activity set. A host country's regulatory, normative, and cultural-cognitive profile will either enhance or inhibit the ability of a foreign manufacturer and local middleman to craft the contractual, ownership, and social elements needed to allow the conduct of channel activities (Grewal and Dharwadkar, 2002). Importantly, a host country's IE may be inadequate or fragile, unable to support IA elements necessary to guarantee economic returns to firms for participating in the channel's activity set. In other words, not every proposed international distribution venture can be conducted; some deals simply cannot be done owing to poor IE support for the type of IA necessary to motivate participation by the appropriate parties (Carson *et al.*, 1999).

The *regulative* element reflects the demands of government bureaucracies, courts, and regulatory bodies to comply with laws and other country-specific rules or requirements (Bello and Briggs, 2009). By imposing constraints and penalties as well as providing inducements (subsidies, tax and tariff concessions, etc.), the legal and regulatory mechanisms in a country strongly influence the feasibility of IA elements to provide the guarantees and economic incentives necessary to secure participation. For example, a US manufacturer may experience little difficulty with its domestic channels in developing contracts with complex clauses detailing cost-plus, take-or-pay, or other pay-for-performance agreements. Ease of contracting in the United States reflects a strong regulative IE due to legal devices such as the Uniform Commercial Code, and because its "independent judiciary strengthens the predictability of the law and makes contract enforcement easier . . . this enlarges the scope of contracting possibilities" (Carson *et al.*, 1999, p. 119). In contrast, some developing countries (*see* EMERGING MARKETS) in Eastern Europe, Asia, and elsewhere lack the comprehensive legal mechanisms, judiciary, and political system necessary to guarantee just and impartial enforcement of contracts. Likewise, because some countries do not strongly enforce property rights, especially for nominal foreign owners, the IA of ownership may not be supported, making certain channel activities infeasible as they are unable to be confidently contracted or owned. Hence, a US firm's market-entry plan for certain foreign markets may be abandoned as impractical because of IEs that are incapable of supporting the required channel activity set.

The *normative* element refers to a society's norms and values that direct behavior through social obligations and expectations, prescribing how actors are to behave, and defining legitimate business practices (Scott, 2001). Consequently, a foreign manufacturer may be unable to motivate the participation of local firms because the required contracting, ownership, and social

elements of the proposed IA are inconsistent with the country's work style and concepts of legitimate business arrangements (Bello, Lohtia, and Sangtani, 2004). For example, domestic Japanese firms tend to respect traditional practices and hold strong loyalties to local partners, making them suspicious and unresponsive to foreign firms making what appear to be highly unconventional and risky business propositions. Owing to mimicking traditional behaviors and internalizing legitimate codes of conduct, potential Japanese partners may object to contractual, ownership, or social aspects of a foreign manufacturer's proposed channel activity set. As a result, the foreign firm may find its market-entry plan for this country impractical because of poor support from normative elements of the IE.

The *cultural-cognitive* element refers to socially constructed, common frames of meaning that achieve compliance by providing people in a country with "prefabricated organizing models and scripts" (Scott, 2001, p. 51), making other types of behavior inconceivable. Cultural control over behavior is habitualized, yielding a collective programming of mind and actions that may or may not support aspects of a proposed IA (*see* BASE OF THE PYRAMID MARKETS: CULTURE INSIGHTS AND MARKETING IMPLICATIONS). Culture is embedded in the perspective of local channel members, impacting their willingness and ability to participate in contracting, ownership, and social aspects of a channel arrangement proposed by a foreign principal. In Japan, culture uniquely supports *keiretsu*-based bonds among local firms that often inhibit a potential Japanese partner from accepting an IA proposed by a firm outside the group. In China, *guanxi* or strong personal connections are often necessary to develop business arrangements. Without the requisite *guanxi*, a foreign manufacturer may find it difficult to craft the necessary IA elements, regardless of the apparent economic and strategic benefits of its proposed channel activity set for the Chinese marketplace (Bello and Briggs, 2009).

To conclude, international marketing channels are one of the most complex and dynamic aspects of global marketing. Cross-border channel arrangements are the result of multifaceted interactions among firms, activity sets, and IAs. Feasible channel designs reflect a compromise among the competing demands of these factors and the constraints associated with the prevailing IE. Clearly, careful analysis and keen managerial judgments are necessary to achieve the most productive balance among the many determinants of international channel design.

Bibliography

Anderson, E. and Gatignon, H. (1986) Models of foreign entry: a transaction cost analysis and propositions. *Journal of International Business Studies*, **17** (3), 1–26.

Bello, D. C. and Briggs, F. (2009) Global channels of distribution, in (M. Kotabe and K. Helsen), *The SAGE Handbook of International Marketing*, Sage Publications, London, pp. 398–412.

Bello, D.C., Lohtia, R., and Sangtani, V. (2004) An institutional analysis of supply chain innovations in global marketing channels. *Industrial Marketing Management*, **33** (1), 57–64.

Bello, D.C. and Zhu, M. (2006) Global marketing and procurement of industrial products: institutional design of interfirm functional tasks. *Industrial Marketing Management*, **35** (5), 545–555.

Carson, S., Devinney, T., Dowling, G., and John, G. (1999) Understanding institutional designs within marketing value systems. *Journal of Marketing*, **63** (special issue), 115–130.

Coughlan, A., Anderson, E., Stern, L., and El-Ansary, A. (2006) *Marketing Channels*, Pearson Education Inc., Upper Saddle River,.

Czinkota, M. and Ronkainen, I.I. (2007) *International Marketing*, Thomson Higher Education, Mason.

Ghosh, M. and John, G. (1999) Governance value analysis and marketing strategy. *Journal of Marketing*, **63** (special issue), 131–145.

Grewal, R. and Dharwadkar, R. (2002) The role of the institutional environment in marketing channels. *Journal of Marketing*, **66** (3), 82–97.

Rindfleisch, A. and Heide, J.B. (1997) Transaction cost analysis: past, present and future applications. *Journal of Marketing*, **61** (4), 30–54.

Scott, R.W. (2001) *Institutions and Organizations*, Sage Publications, Thousands Oaks.

Williamson, O.E. (1996) *The Mechanisms of Governance*, The Free Press, New York.

Zhao, H., Luo, Y., and Suh, T. (2004) Transaction cost determinants and ownership-based entry mode choice: a meta-analytic review. *Journal of International Business Studies*, **35** (6), 524–544.

international retailing

Brenda Sternquist

The 1990s has been an era of consolidation and internationalization for major food retailers (Sternquist, 2007). CIES, the Food Industry Forum, conducts an annual membership survey. Internationalization of retailing was one of the top three issues for 1998 and continues to be a major industry interest. We know little about globalization of the retail industry. We are used to analyzing globalization in manufacturing, but retailers have stayed home, until now.

Manufacturers are interested in retail internationalization. They want to know how to deal with their retail customers who buy, not on a country to country basis, but on a regional or global basis. Global/Multinational retailers have the capacity to bring about regional, and possibly world, price convergence. Global-NetXchange (GNX) is an international Internet sourcing system sponsored by Sears, Carrefour, Kroger, Metro, Sainsbury, Coles Myer, and Pinault-Printemps-Redoute. Sainsbury has used this system's reverse auction reducing the bidding process from 6 weeks to just a few hours. Sainsbury anticipates purchasing 75% of their goods through GNX.

There has been very little theoretical development leading to an explanation of how, why, and where retailers internationalize. Internationalization theories that are developed to explain manufacturer's internationalization have limits when applied to retailing. Most theories of internationalization begin with the assumption that the company has an export option but retailers do not have an export option. Since the retail function is one that requires sale to an ultimate customer they must have a physical presence in the country to be a retailer. This presence might take the form of a website in the case of an Internet retailer, but it is a presence.

Two major problems have limited the development of a comprehensive theory of retail internationalization. The first limit is that most researchers have assumed that internationalization is the same in all types of retailers. Salmon and Tordjman (1989) were the first to identify two different types of retailers that they termed *global* and *multinational*. However, other than Sternquist (1997a) major researchers have not followed this designation. Instead, all types of retailers are put together for analysis. There are at least four different types of retailer internationalization (investment, pure franchisor, global, and multinational) each with a different expansion pattern, partly because of the way the company is organized, but there are other factors that also influence internationalization choice. Therefore, differentiation of retailer types helps to distinguish how different retailers internationalize. The second major problem is that most researchers have observed only retailer's successful international moves without analyzing the actions that are not successful. It is essential to analyze the internationalization failures in order to fully develop a model of successful retail internationalization. Failures are important because companies are more likely to analyze failures rather than success. Also, failures represent the definitive answer about whether an expansion pattern is correct.

Types of Retail Internationalization

Investment. Investment refers to purchasing an existing well-managed company and allowing management to continue operations. This is a dormant type of management. This is the strategy used by Delhaize, Ahold, and Sainsbury when they first entered the United States. Investment also appears to be the strategy used by Wal-Mart in their purchase of ASDA (UK). Dave Ferguson, president and CEO of Wal-Mart Europe said, "We've got a company that we were smart enough not to try and change" (ASDA, 2001). Companies using investment strategy are looking for (i) higher than domestic return on investment, (ii) a safe foreign international investment environment and/or, (iii) access to retail know-how. The motivation for purchasing companies in other countries rather than the home market is to diversify the company's investment. Delhaize purchased Food Lion in the United States at the height of the cold war. They wanted the chance to put company resources in a more stable geo-political environment.

Internationalization through investment leads to little knowledge transfer, and therefore it does not necessarily lead to more active models of

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

management in the company's internationalization. Investment is also an expensive internationalization strategy because the company is buying a well-managed company, rather than a company in trouble, and are therefore paying top dollar for this investment, which makes the return on the international investment difficult to recoup. Ahold sought the "gem" retailers in each country they entered. If they could not purchase the country gem they would not enter the market. We can see from the recent disposal of Ahold acquisitions in South America and Asia that the investment strategy is difficult to pull off.

Pure franchisors. Another form of dormant management is a pure franchisor. These companies use franchising exclusively in their international expansion. Pure franchisors typically focus on master franchising networks given to individuals on a countrywide basis. Pure franchisors typically expand internationally for opportunistic reasons. Opportunistic means that someone approaches the company with the idea for a franchise, rather than the company developing an interest in the country for strategic purposes. For instance, the McDonald's Japan master franchise was given to an entrepreneurial handbag/shoe salesman having an account at Mitsukoshi. He convinced the McDonald's corporation that their burgers and fries would be a hit in Japan. At that time McDonald's was not interested in opening company stores in Japan. For master franchisors the identification of a great person to develop the network is more important than identifying the best country to enter. For this reason, when one views the expansion pattern of master franchisors, it is of a random variety of countries rather than a stages type of expansion where a company sequentially enters culturally similar countries.

Both investment and pure franchisor internationalization are based on dormant management. Because management is dormant there is little knowledge transfer. Also, from a retail strategy standpoint, neither method is a strategic retail decision. Both of these options are financial decisions that do not lead to further international expansion in a predictable way. It is important to identify these two motivations for international expansion because retailers using these approaches need to be excluded from modeling the internationalization of retailers as they are not strategic and rational, but rather opportunistic and random. Events that are random cannot be successfully modeled. In other words, you cannot take their experiences into consideration when developing propositions that will explain internationalization.

Two types of retailers whose moves can be considered strategic are focused on here. The terms initiated by Salmon and Tordjman (Sternquist, 2007), namely, global and multinational retailers, are used. The terms *global* and *multinational* do not indicate their degree of internationalization; instead they refer to the concept of standardization (global) versus adaptation (multinational). There is controversy about these terms. In the marketing literature, multinational firms are often referred to as multilocal. But this concept does not convey the idea that multinational retailers learn from each international movement and this knowledge is important in determining future international actions. Not every retailer can be classified as a totally global or a totally multinational retailer. Most retailers have some characteristics that cross over, but we can see some patterns when we use this designation.

Pellegrini (1994) presents an important distinction to explain paths for growth. He identifies the search for growth related to (i) an attempt to extend the application of the firm's proprietary know-how to get the benefits of internationalization; this is called *rents* in the management literature; (ii) an attempt to optimize the scale of operations (economies of scale); or (iii) an attempt to expand the mix of operations (economies of scope) to reduce costs and increase efficiency. In the next section the paths for growth for global and multinational retailers are considered.

Global. Global retailers are centralized, standardized, and generally small format retailers. These retailers are often vertically integrated, frequently focusing on private label or exclusive merchandise. Examples of global retailers are Zara, Mango, and Gap. They do not change their retail offering much when they enter foreign markets, but they look for a universal, global market segment that will accept their retail offering. Since these retailers are relatively

small in scale and centralized, it is possible to develop business format franchises for the system. Because they have a standard format, they can be replicated very rapidly in new locations.

Global retailers begin their internationalization not to countries, but to world class cities where they expect to find a segment of consumers that are indistinguishable throughout the world. They do not need to change their offering because their consumer is fundamentally the same wherever they go.

Global retailers, particularly fast food retailers have been discussed as new forms of global imperialism. Global imperialism means that the retailers alter material culture wherever they go. Authors such as Ronald Steel and Thomas Friedman argue that McDonald's and other manifestations of global culture serve the interests of middle classes that are emerging in autocratic, undemocratic societies.

Global retailers' motivation for international expansion is to exploit a unique concept before others have a chance to do this. Global retailers therefore may expand to other countries without saturating their home market. Their quest is not a simple quest for growth but a need to colonize their concept.

Colonizing a concept means that one expands very rapidly before others can mimic one's idea. Although all global retailers expand internationally more rapidly than multinational retailers, there is a group of exceptionally fast internationalizers. These companies expand internationally nearly from their inception. In the marketing literature, Knight and Cavusgil (1996) term these companies *Born Global*. The Born Global literature focuses on learning in networks to explain the quick international expansion. The early movers in retailing do not depend on learning networks, but instead focus on colonizing the concept. Being first with the idea gives significant first-mover advantages. They are not changing their offering, so there is no real need to learn about the international environment.

Multinational. The second type, a multinational retailer, is decentralized and adapts its product offering to the culture it is serving. It is usually a large-scale format, which would be difficult or impossible to develop into a business format franchise. If this type of retailer wishes to use a low-cost alternative for international expansion, they do not have the option of franchising and will have to use licensing instead. Wal-Mart's Supercenters, warehouse clubs such as PriceClub, and supermarkets such as Tesco are multinational retailers. These retailers adapt to the cultures they are entering. They generally source merchandise locally and try to hire and train management from the local area.

We can best understand multinational retailer's internationalization as a strategy for growth. These retailers do not generally internationalize before they have saturated their home market. This saturation may be because of natural market expansion or, as we more often find in Europe, government restriction on large-scale store expansion. This can explain why European food retailers have been the most aggressive in cross-border expansion.

Some of the first multinational retailers to expand internationally were the Japanese department stores and General Merchandise Stores (GMS). These companies moved into Hong Kong and Taiwan to avoid the competitive environment of Japan where large-scale store laws prevented their expansion at home. These early international moves were defensive. The first Japanese department stores to internationalize such as Yaohan were weak in their home market. In the end, most of the early Japanese department store retailers' international expansion failed. A second wave of Japanese department stores expanded to Hong Kong and Taiwan with the intention of learning how to do business in the Chinese business environments. Companies such as Isetan were successful in using Hong Kong and Taiwan to launch operations in China.

Most large-scale European retailers that internationalized were food retailers (Sternquist, 1997b). They did not generally enter foreign countries until they had saturated the home market. Growth needs fueled their international expansion. In this case, unlike the Japanese early expander, international expansion was part of their strategic growth.

Since multinational retailers have a large-scale format, they do not have a franchising option. Franchising requires that the business format

be such that it can be standardized, essentially codifying the knowledge, which can then be transferred to foreign outlets. Multinational retailers who want to use a low-risk alternative must use licensing rather than franchising. Licensing is a risky move for a company because they have little control over what the licensee will do with the name.

Theoretical Explanations for Internationalization

Stages theory focuses on the pattern of internationalization. According to stages theory, a company will initially expand to countries that are most similar to their home market. Then with experience they will expand to countries that are less similar. Dunning's (1981) eclectic theory focuses on ownership, location, and internalization factors. Institutional theory explains the influences of macroenvironment and microenvironment on a company's decision to internationalize. The final conceptual link is risk theory. Retailers evaluate foreign markets according to their perceived level of risk. Much of a retailer's in-country investment is for inventory and physical facilities, an investment that is not liquid, and therefore not salvageable in the short run if a government determines that they should not operate there anymore, or if they simply cannot survive.

The export stages literature views internationalization as a step-by-step process where higher level stages represent a higher level of international involvement than lower stages (Cavusgil, 1982, 1984). This paradigm relies on characteristics of management to delineate the stages they are ready to assume. However, the stages available for retailers are different from those available to manufacturers. Retailers do not have an export option. The stages for retailers are based more on risk and knowledge. Multinational retailers will begin their international expansion in countries that are culturally similar to the home country. They will develop this area or region and then make a jump to another geographic area which is not culturally similar, and develop area expertise through local expansion. As they gain experience in each country or region, they move into another area. Eroglu (1992) provides the basis for predicting which firms will seek international expansion because of organizational characteristics such as size, experience, and international orientation. These organizational characteristics are directly related to the propensity for international expansion.

Retailers make mistakes in their internationalization efforts. A failure can be very valuable for a retailer's long-term success because failures are analyzed more than successes. Analysis leads to learning. Carrefour is a good example of a company that withdraws from markets in which it is not successful. As a result of this trial learning, the company has compiled a vast portfolio of information about how to do business in other countries. Carrefour's most recent move was to acquire GB, a Belgium firm. Belgium was the first country Carrefour entered (1969). They withdrew after their initial entry. However, after over thirty years of international experience they are ready to reenter through acquisition.

Researchers have developed most international business models from the manufacturing perspective. The two major international business research paradigms are the behavioral school and the export stages view. The behavioral school considers internationalization as an incremental process of strategic growth based on a firm's general and experiential market knowledge and resources. Increased market knowledge enables the firm to have higher international involvement. This view assumes that knowledge and experience accumulates in a firms' internationalization growth and allows the firm to control risks associated with internationalization. The behavioral internationalization paradigm has been used as a model for international retail involvement. Results confirm that strategic management characteristics, competitive advantage related to retail concept and logistics and retailer's size are related to retail internationalization (Vida, Reardon, and Fairhurst, 2000).

Both these models assume that international activities are at higher risk than domestic expansion, the internationalization process is sequential (lower involvement leading to higher involvement), and internal firm factors influence international involvement rather than macro-environmental factors.

Retailing differs from manufacturing internationalization. Retailers often do not use sophisticated decision models before internationalizing.

They are more likely to take an international move and then assess the impact of the move. Yip, Biscarri, and Monti (2000) observed this ad hoc approach when they studied the internationalization of small and medium-sized firms. Retailers also consider country economic characteristics more than manufacturers do, because they must have a physical presence in the country. Therefore, a theory to explain retailer's international expansion should include: (i) the exits, as well as the entry, into foreign markets and (ii) within-country competitive, economic, and development factors.

Dunning's eclectic theory Sternquist (1997b) focuses on ownership (O), location (L), and internalization (I) factors that influence a company's decision to internationalize. In addition, institutional theory helps explain how a retailer's macroenvironment – the host country and home country – and microenvironment – the retailer's past experience and competitors' move – affect the retailer's internationalization decision.

Eclectic theory.

Ownership advantages (O). Ownership advantages include innovative or unique products or processes the company can use to obtain market power. This grouping includes asset-based and transaction-based advantages. Asset-based advantages refer to tangible items such as patents or unique products. Global retailers generally have significant asset-based ownership advantages, since they often focus on private label merchandise. Private label can be considered as an important ownership advantage. However, private label does not necessarily translate well when a company internationalizes. Retailers such as Zara and Mango are defined by their private label concept. In this case, the private label is an asset in international expansion. Also, because private label is easily expanded to international settings it makes the asset transportable. In contrast, multinational retailers are generally mass merchandisers and for these retailers private labels do not transfer well in an international setting. A good example of this is Marks & Spencer, a UK company that focuses on general merchandise and food offerings. Although they had a very successful start at internationalization in selling their products in Hong Kong (a former British colony) and other communities where there was a strong affiliation to the Marks & Spencer concept, this success diminished when they had less of an expatriate clientele. As a result they have had to back out of many markets where they thought they would be successful. Gielens and Dekimpe (2001) found in their study of European food retailers that private label was not related to sales and efficiency. The bottom line is that very targeted retailers such as Mango and Zara are looking for a distinct segment of consumers who like their product throughout the world, but large-scale retailers have a difficult time using their private label effectively because the mass population are not familiar with, and do not accept, private label merchandise from a retailer they do not know.

Transaction-based advantages come about because of the way things are done. Customer service or centralized buying are examples of transaction-based advantages. Transaction-based knowledge is highly tacit; transferring it to others is difficult and is characterized through learning by doing (tacit learning). Transaction-based ownership advantages are difficult to transfer to another country. This both protects this type of ownership advantage from being copied, but also makes it difficult to use franchising to transfer the knowledge.

Transaction-based advantages include knowledge about how to internationalize. Carrefour and Metro are two food retailers with long-standing experience in emerging markets. This has given them ownership advantages in three major areas. The first is in people. Carrefour's years of international experience have given them a pool of experienced managers. These managers can be used to open new stores throughout the world. Carrefour has decentralized stores, giving managers extensive decision-making experience. Recently, Carrefour has been moving toward greater centralization to gain economies of scale in purchasing. The second area advantage is the hypermarket/warehouse club concept. This concept is particularly well adapted to emerging markets, where consumer purchasing power is growing rapidly from low initial levels. Also, large store formats benefit from being in areas where land is reasonable,

available, and not heavily regulated by the government. The third area of ownership advantage is their negotiating position with suppliers. Carrefour's international expansion is a textbook example of how a multinational retailer internationalizes. Carrefour's early international success came about because of their "first mover" advantage. Their first internationalization efforts were in Northern European countries. These movements were not successful and the company retreated. Over time the company began to expand to less developed countries such as Spain, where they had a great deal of cultural similarity. They then expanded to South America, where they were able to capitalize on their knowledge of doing business in Latin cultures. As other multinational food retailers entered the markets and competition increased, they have begun to struggle (Evans and Mavondo, 2002). In 2005, Carrefour withdrew from Mexico, forced out by Wal-Mart's initiated price wars.

Interviews with retailers revealed that they internationalized because they had this "wall of cash" that they had to do something with. They had been successful and now they had to reinvest their earnings. Also, many retailers said that the investment community expected them to internationalize, and would look at their company unfavorably if they did not move outside their home borders. This means that stock market analysts would not suggest investing in their company if they did not expand. So, some international expansion is related to satisfying the investment community, a source of revenue for publicly held companies.

Locational advantages (L). Locational advantages relate to how suitable the host country is with respect to the firm's strategies. In other words, how well does the country provide a comfortable place for the retailer to reside? Pellegrini's (1994) work identifies issues relevant to retailing.

Cultural proximity: People in some countries share similar patterns of life. Midwesterners in the United States are more like Midwesterners in Canada than US residents on either coast. Europe has three bands of culturally similar groups divided geographically by horizontal bands, rather than distinct country groupings. French hypermarkets dominate the Spanish market. German department stores have expanded rapidly into what had been East Germany. These are examples of retailers expanding into culturally similar environments. Cultural proximity is more important for multinational retailers, and becomes insignificant when retail concepts involve narrowly defined consumer markets, namely, global retailers.

Evans and Mavondo (2002) considered the relationship between psychic distance and organizational performance. Psychic distance is conceptually the opposite of cultural proximity. They found that psychic distance has a positive effect on organizational performance. Their explanation for this finding is that retailers who perceive a difference in culture between the home and the host country will make adaptations while those who perceive no distance will not make adaptations. This finding is an example of where the type of retailer is likely confounding the results. If these authors had separated retailers into global versus multinational they would likely find that psychic distance is not important to global retailers because they are selling to essentially the same customers throughout the world. Global retailers will enter world cities, selling essentially the same thing they sold at home. Multinational retailers will enter countries initially that have the smallest cultural distance to the home market, but over time they will expand to countries that are culturally distant.

Market size: There are both push and pull factors related to market size. Push factors mean that the retailer's growth in the home market is limited and they need to seek other environments. Saturation in the home country is an impetus for foreign direct investment by multinational retailers. Ample space for expansion must be available, particularly if the firm needs to reach a certain size to exploit economies of scale. Legal restrictions governing growth in the home country can be considered a motive for companies moving to foreign locations. France, Belgium, Germany, and Spain have stringent requirements that

limit large-scale retailers' growth, pushing retailers to expand outside national borders.

The French government passed a law in 1996 that set much more restrictive limits on new store openings. The law has essentially put an end to new hypermarket openings. Hard discount retailing, the small-scale version of warehouse clubs, will also be affected by the law. Hard discount stores are typically 900 m^2 in size. The new legislation lowers the minimum store size requiring authorization to 300 m^2. This limits further growth even for hard discounters.

Market pull factors are those that make a foreign market look attractive. Multinational retailers will seek countries that are less developed than their home market. Markets with high growth in income can provide retailers with double-digit growth. Global retailers are not affected by the level of economic development in the country. They seek world capitals where a homogeneous consumer resides.

Multinational retailers who enter less developed countries are also expected to alter their product offering so that food becomes a larger percent of sales. Consumers in less developed countries spend a higher percentage of their total disposable income on food. Wal-Mart sold very little food in its stores until they entered Mexico. In Mexico, their joint venture partner, Cifra was the country's major food retailer. Their joint venture businesses had a heavy focus on food. This joint venture has subsequently been terminated.

Competitors' moves: A first-mover advantage may be lost if competitors enter a foreign market. Competitors may secure prime retail locations and block out other firms. Within any major city, there are only a few prime retail sites. The larger the retailer considering international expansion, the more crucial it is to be an early mover. Like a strategic game of chess, the early moves of a large retailer sets the stage for a long-term competitive advantage. Competitor's moves have little impact on global retailers.

Geographic proximity: Expanding closer to home reduces costs related to transportation and corporate communication. Geographic proximity is important for retailers selling private labels produced in a central location. Geographic proximity is less important for decentralized companies because they are allowed to operate as independent units, generally sourcing from within the country. Even before the peso devaluation, Wal-Mart Mexico purchased 80–90% of their products from Mexican sources.

Low-cost land and labor: Carrefour has a policy of owning their stores. When Carrefour opens in a new commercial zone, it attracts the best tenants to the hypermarket complex. This gives them a strong economic justification to own their own outlets, and the whole shopping center, if possible. Only in countries where the cost of commercial land is exorbitantly high, or where, for legal reasons, they cannot obtain title, does Carrefour agree to be a tenant. This was the case in Taiwan until 1996 when the government reduced restrictions on foreign ownership of land. Even in China, where land cannot actually be owned, but only obtained through a long-term lease, Wal-Mart builds its own stores.

Internalization (I): The greater a company's ownership assets, the more important it is to protect these assets by guarding company secrets. Internalization entails keeping the information within the company. In retail franchising, the company sells or leases ownership assets to other firms. Therefore, franchising is a particularly dangerous idea for retailers with strong transaction-based ownership advantages. Retailing innovations are difficult to defend from imitators. Competitors can freely copy them because there are no patents on retail know-how; therefore, to maintain a competitive advantage, the retail company needs to internalize its innovations. The only way for a retail company to keep its operating secrets is to open wholly owned subsidiaries in various countries. This is too expensive for many companies. They are forced to franchise their retail innovations to expand rapidly. However, over time they may regret this decision because they will have lost organic growth opportunities in that area. There is one instance where franchising is

actually a smart strategic move, that is, if you know that your retail format will be copied rapidly and you have to franchise to be able to move quickly enough.

A primary attraction for the use of franchising in overseas markets might be rapid expansion. For some companies, foreign markets may be viewed as providing favorable opportunities for growth, regardless of the level of development of the domestic market. Many retailers, for example, *Body Shop*, *Benetton*, *Marks & Spencer*, and *Mango*, have increasingly adopted franchising as an expansion method. This tendency is consistent with the premise that the global retailer with an innovative concept expands rapidly to fully exploit the ownership asset because expansion is rapidly accomplished through a franchising system. Many retail firms use other strategies to achieve rapid growth objectives. With a highly innovative retail and brand concept, and a primary growth method through wholly owned subsidiaries, *Zara* almost instantly expanded into 30 countries, adding nearly 250 stores from 2001 to 2003. Also, *GAP* with well-defined concepts has used direct investment, rather than franchising, and achieved rapid expansion (Park and Sternquist, 2008). Given the existence of global retailers using the wholly owned entry mode to expand rapidly, there may be some intervening factor that influences a global retailer's mode choice of entry with regard to rapid expansion.

There are two theoretical explanations for franchising, namely, agency theory and resource based theory. Agency theory suggests that companies franchise because monitoring of the foreign subsidiary is very difficult, and therefore making the foreign entity an owner manager will reduce the tendency to cheat. Resource-based theory suggests that firms franchise because they cannot afford to expand on their own. Although we cannot predict which theory explains a firm's decision to franchise when they begin the internationalization experience, over time we can have an idea of the correct explanation. If the company begins to reacquire their franchise outlets on amassing resources, then this would lend support to the resource-based theory. In other words, if they find themselves with money and they use this money to buy back franchisees then we can conclude that it is resource-based theory that is in play.

Joint ventures are when two companies join together to create a new entity. For instance, when Wal-Mart went to Mexico, they entered into a joint venture with Cifra to open super-centers called Bogata. In a joint venture partners share information and knowledge, which means that joint ventures offer less protection of secrets than wholly owned subsidiaries. Sometimes joint ventures are needed when entering a different cultural environment or because of government regulations. Many developing countries require that domestic ownership have controlling interest of multinational companies. International retail joint ventures nearly always involve a multinational retailer. Generally, the foreign retail joint ventures involve a foreign company that needs to learn the culture and business practices in a new country. They select a joint venture partner in the country that will provide them with this type of information. However, once the foreign retailer learns how to do business in the foreign country they are likely to acquire the newly formed enterprise. Therefore, retail foreign joint ventures are posited to be temporary in duration. The concept that foreign joint ventures are only useful until the foreign company has learned the culture is interesting.

Other Theoretical Explanations for Internationalization

Institutional theory. Institutional theory emphasizes that a retailer's institutional environment influences its decisions and behaviors (Huang and Sternquist, 2007). Here, the institutional environment not only consists of regulatory structures, laws, rules, cultural beliefs, norms, and habits but also customers, suppliers, and retailers itself. Retailers' international expansion leads to the transfer of retail management technology across boundaries. External environment factors such as different cultural, legal, political, social, and economic factors affect retailers before they make decisions as to where, when, and how to expand into another country. Half of the top 50 companies in the S&P 500 were established within the past 20 years, yet they began to internationalize only within the past 10 years. These companies include retailers such

as Home Depot and Best Buy. As retailers increasingly consider international expansion as a viable growth option, entry location, entry timing, and *entry* mode remain as the cornerstones of retailers' market entry strategy. For retailers, franchising (or licensing), entering into a joint venture and setting up a wholly owned subsidiary are most distinct modes of entry into a foreign market. Different entry modes represent different levels of *resource commitment*, and control over the foreign operation.

Scott and Christensen, (1995) classified three kinds of institutional systems – *regulative*, *normative*, and *cognitive* systems representing legal, social, and psychological elements and the focus here is on the influences these three dimensions on retailers' decision making and behaviors of international expansion.

Regulative dimension. Institutional theory emphasizes legal and political factors specific to a country. The home and the host country's legal regulations represent the strongest environmental pressures faced by retailers. For example, the Large-scale Retail Store Law (LSRS) in Japan not only significantly impeded the expansion of foreign retailers such as Toys "R" Us (Scott and Christensen, 1995), but also pushed Japanese retailers such as Isetan to expand internationally.

A country with a strong rule of law is defined as one having sound political institutions, a strong court system, and provisions for orderly succession of power. A strong rule of law represents freedom of transacting, security of property rights, and transparency of government and legal process. The rule of law is a double-edged sword. On the one hand, rules establish a stable structure to reduce uncertainty about what protection retailers can expect from the host country's legal system. On the other hand, foreign operation restrictions can be the barriers to retail firms.

Countries with a weak governance infrastructure have to improve to be able to attract foreign investment. Researchers have found that the improved legal environment is a significant determinant of the timing of entry. Retailers are more likely to establish wholly owned subsidiaries in transitional economies that have progressed furthest in institutional reform. However, if the rule of law gets more and more restrictive, such as Japanese's LSRS law, retailers may lose interest in expanding or prefer entering using franchising or licensing.

Normative dimension. The normative dimension of the institutional theory gives priority to norms and cultural influence. For a retailer who sells consumer goods overseas, the cultural difference between consumers from the home and the host country will probably be one of the most important factors to be taken into consideration when determining the form of market entry. Toys "R" Us expands with some adaptation to different markets. Though the overseas stores are formatted similarly to those in the United States, its merchandising selections are different. Twenty percent of its merchandising assortment is chosen for local consumer interests. For example, the Japanese market prefers porcelain dolls whereas Germans prefers wooden toys.

The cultural distance between the retailer's home country and the host country affects the choice of foreign expansion form. Many researchers support the theory that greater cultural distance results in low control entry modes such as franchising and licensing. In the case of high cultural distance, retailers may perceive high risk of entering a foreign market and feel high pressure of serving different customers. Accordingly, they may prefer not to enter the country or delay the entry. However, if the benefits to be reaped in the foreign market are high enough, retailers may still wish to enter into the country. In order to "fit" into the new environment, retailers may choose adaptation. Wal-Mart did not initially adapt its retail format in Argentina to the local culture. However, the retailer learned valuable lessons there that helped its subsequent operations in similar situations.

Adaptation is a long-term and accumulative process. Retailers may seek local partners to accelerate the process. On the contrary, retailers are more likely to enter countries with cultures that are similar to the home market before entering countries with less similar cultures. In this case, retailers may choose wholly owned subsidiary because a high level of understanding of norms and values has already been shared; therefore, local partners are less needed.

Cognitive dimension. The cognitive dimension of the institutional theory explains why a global retailer such as The Body Shop always uses the same entry mode during international expansion. From a cognitive perspective, organizations possess habits and inertia. As Porter (1990) notes: "Firms would rather not change . . . Past approaches become institutionalized in procedures and management controls . . . " Retailers tend to use entry modes consistently, especially when the situations are similar to the past. The primary form that the internal institutional environment can influence entry mode choice is through habit, which means that once a practice or decision has been chosen and implemented, the likelihood of alternatives being considered and used in future decisions will be reduced. Habits and inertia preclude rational changes.

Further, retailers can learn not only from their own experience but also from the experience of others, such as competitors. Strategic choice theories suggest that imitation can be a strategic response to competitor activities, whereby second-movers take advantage of the fact that the risk associated with new situation has been absorbed by the first-movers. In other words, copycats can be successful too. A retailer could decide which country to enter by following other retailers. For example, China has been one of the most popular investment destinations in the world. For several years in the 1990s, China was the second largest recipient of FDI worldwide. One motivation for foreign firms investing in China is to meet their competitors' move to China and/or forestall or reduce competition from other firms' earning markets in the same geographical area. When considering a totally unknown foreign area where minor similarities could be found in terms of previous practices, retailers will naturally avoid or try to postpone the consideration first. Nevertheless, if the retailer decides to expand into this new area and has no past experience to rely on, the retailer may resort to other retailers' experience, that is, to mimic others expansion behaviors.

Risk theory. The final theory is related to risk. Multinational retailers will begin their international expansion in countries that are culturally similar to the home country. As they gain experience in each country or region, they move into another area. In Eroglu's (Cavusgil, 1984) conceptual model, she provides the basis for predicting which firms will seek international expansion because of organizational characteristics such as size, experience, and international orientation. These characteristics reduce risk or increase knowledge which is needed to handle risk.

Risk theory also relates to how much investment a company puts into international expansion. Global retailers who seek a less risky alternative can select to franchise. Multinational retailers who seek a less risky alternative will have to license. Franchising and licensing both require the relinquishing of company assets.

Summary and Conclusions

This article provides the conceptual overview for understanding international retailing. The study of international retailing and retail internationalization is very different from the study of international marketing and market internationalization. Retailing is geographically based. Laws regulating international trade do not affect international retailing, as the retailers can source in each domestic market. Instead, retailers need to be aware of intracountry business laws and regulations.

Retailers must assess their company's strengths and the relative attractiveness of continued domestic growth or international growth. They must determine whether it is important to keep company knowledge to themselves and use it for their continued international expansion.

Retailers use the macroenvironment – economic, competitive, technological, social, and governmental – and the microenvironment– competitors, past experience– to make decisions about their retail offering. This decision is twofold: first, should they enter the foreign market? And, if they decide to enter the market, should they use a standard retail format (global), or should they adapt their retail offering (multinational)? The issues are who should go, where to go, and how to go. Such issues are also related to the decision to rely on centrally versus decentrally managed operations.

The macroenvironment influences a country's locational advantage. A company's ownership

advantages will influence whether it will internalize its secrets. This decision will, in turn, influence whether the company uses a standard format or a format that is individualized for each country.

Bibliography

ASDA (2001) A model acquisition. *Chain Store Age Executive*. **77** (6) 58.

Cavusgil, T.S. (1982) Some observations on the relevance of critical variables for internationalization stages, in *Export Management: An International Context* (eds M.R.Czinkota and G. Tesar), Praeger Publishers, New York, pp. 55–62.

Cavusgil, T.S. (1984) Organizational characteristics associated with export activity. *Journal of Management Studies*, **21** (1), 3–50.

Dunning, J.H. (1981) *International Production and the Multinational Enterprise*, Allen & Unwin, London.

Eroglu, S. (1992) The internationalisation process of franchise systems: a conceptual model. *International Marketing Review*, **9** (5), 11–39.

Evans, J. and Mavondo, F. (2002) Psychic distance and organizational performance: An empirical examination of international retailing operations. *Journal of International Business Studies*, **33** (3), 515–532.

Gielens, K. and Dekimpe, M. (2001) Do international entry decisions of retail chains matter in the long run? *International Journal of Research in Marketing*, **18**, 235–259.

Huang, Y. and Sternquist, B. (2007) Retailers' foreign market entry decisions: an institutional perspective. *International Business Review*, **16**, 613–629.

Knight, G. and Cavusgil, T. (1996) The born global firm: a challenge to traditional internationalization theory. *Advances in International Marketing*, **8**, 11–26.

Park, Y. and Sternquist, B. (2008) The global retailer's strategic proposition and choice of entry mode. *International Journal of Retail and Distribution Management*, **36** (4), 281–299.

Pellegrini, L. (1994) Alternatives for growth and internationalization in retailing. *The International Review of Retail, Distribution and Consumer Research*, **4** (2), 121–148.

Porter, M.E. (1990) *Competitive Advantage of Nations*, Free Press, New York.

Salmon, W. and Tordjman, A. (1989) The internationalization of retailing. *International Journal of Retailing*, **4** (2), 3–16.

Scott, W. R. and Christensen, S. (eds) (1995) *The Institutional Construction of Organizations*, Sage, Thousand Oaks.

Sternquist, B. (1997a) A conceptual model of strategic international retail expansion. *International Journal of Retail and Distribution Management*, **8** (25), 262–268.

Sternquist, B. (1997b) Internationalization of Japanese department stores. *International Journal of Commerce and Management*, Special Issue on Global Retailing. (7) **1**, 57–73.

Sternquist, B. (1998) *International Retailing*, Fairchild Press, New York.

Sternquist, B. (2007) *International Retailing*, 2nd edn, Fairchild Books, New York.

Vida, I., Reardon, J., and Fairhurst, A. (2000) Determinants of international retail involvement: the case of large U.S. retail chains, *Journal of International Marketing*, **8** (4): 37–60.

Watson, J. (2000) China's bigMac attack. *Foreign Affairs*. May/June, 120–134.

Yip, G.S., Biscarri, G., and Monti, J.A. (2000) The role of the internationalization process in the performance of newly internationalizing firms. *Journal of International Marketing*, **8** (3), 10–35.

designing a global supply chain: opportunities and challenges

David J. Closs

Globalization (*see* GLOBAL CONSUMERISM AND CONSUMPTION) offers many opportunities and challenges for logistics and supply chain operations and strategies. The opportunities include increasing markets and a wider range of manufacturing alternatives with varying absolute and comparative human and material resource advantages (*see* COMPETITIVE ADVANTAGE: ITS SOURCES AND THE SEARCH FOR VALUE). Some regions of the world can offer significant economies of scale because of their competitive wage scales, while other regions offer significant flexibility because of their expertise. The challenges encountered when taking advantage of these benefits include more demanding logistics operating environments, security considerations, and more complex total cost analyses (Bowersox, Closs and Cooper, 2010).

Global Economies

Regardless of size, most firms today include some dimension of global operations. They have materials or products that are sourced globally or global customers wanting to purchase from them (*see* GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES). In many cases, firms are involved in both global sourcing and delivery. Table 1 lists the primary rationale for firms to globalize. While many believe that the primary motivation for shifting manufacturing and supply chain operations is low-cost resources and labor, the rationale in many cases is one of the others cited in the table. For example, although it is believed that many firms moved production to Asia and India to access lower production wages, their primary motivation was often to obtain access to rapidly growing markets. While many have taken advantage of rapidly increasing demand, the increased demand for employees in production is driving up wage rates much more rapidly than in the developed world. As a result, the wage differential is declining and firms are beginning to look for the next low-cost production sources which today include Vietnam and Cambodia. Firms are also beginning to develop supply chain value-added capabilities in South America and Africa to take advantage of market proximity to the developed world and low-cost production.

Global supply chain operations either directly or indirectly is becoming the norm for most firms. Global sourcing and marketing offers many opportunities to enhance the firm's performance particularly in terms of revenue, volume, and market share (*see* GLOBAL SOURCING STRATEGY: AN EVOLUTION; EMERGING MARKETS). The following section describes some of the strategies firms employ to achieve these benefits along with some of the challenges that firms are likely to encounter.

Global Supply Chain Integration

Just as an effective logistics system is important for domestic supply chain integration, it is absolutely essential for successful global sourcing, manufacturing, and marketing. Domestic logistics focuses on performing movement and storage activities to support supply chain integration in a relatively controlled and consistent environment. Global logistics must support operations in a variety of different national, political, and economic settings while also dealing with increased uncertainty associated with the distance, demand, diversity, and documentation of international commerce (*see* FORCES AFFECTING GLOBAL INTEGRATION AND GLOBAL MARKETING).

The operating challenges of global logistics systems vary substantially in each major global region. The North American logistics challenge is one of an open geography with extensive and relatively integrated transportation options and limited need for cross-border documentation. The European logistician, in contrast, is confronted by relatively compact geography involving numerous political, cultural, regulatory, and language situations. The European infrastructure is also quite congested because of population density and the fact that many of the roads date back centuries. The Pacific Rim logistician faces an island-based environment with relatively poor infrastructure, requiring extensive water and air shipments to transcend vast distances. These different

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Rationale for globalization.

Objective	*Rationale*
Increase revenue	• Open up more markets • Expand beyond competitors • Obtain accessibility to markets that limit access without local operations
Achieve economies of scale	• Take advantage of available production capacity allowing sales of more volume
Reduce direct cost	• Take advantage of lower labor rates or real estate expense • Reduce energy requirements by reducing distance or changing transportation mode • Take advantage of differences in production requirements
Advance technology	• Obtain access to advanced technology that may not be available from current locations due to historical investments • Obtain access to specialized expertise or language skills
Reduce firm's global tax liability	• Obtain local or regional tax benefits related to property, inventory, or income • Obtain reductions in value-added-taxes due to localized production or other value-added services (i.e., packaging, inventory management, customization)
Reduce market access uncertainty	• Source product from location that involves less transportation uncertainty • Source product from location that involves fewer security constraints
Enhance sustainability	• Source products or other resources (including human resources) from locations that have ongoing availability of materials and expertise such as energy or trained workers

characteristics require that firms having global operations develop and maintain a wide variety of capabilities and expertise.

In the past, an enterprise could survive by operating with unique North American, European, or Pacific Rim business strategies. While it was easier to create and operate unique regional strategies, the resulting duplication often resulted in loss of economies of scale and poor asset utilization. While regionalization remains viable for some firms, those desiring to grow and prosper must face the challenges of designing and operating a globally integrated enterprise. Strategic business initiatives must change as a firm and its supply chain become progressively more global.

Stages of International Development

The continuum of global trade development ranges from the lack of an international strategy to the concept of a globally integrated enterprise. The following discussion describes conceptual and managerial implications of strategic development (*see* STAGES OF MARKET DEVELOPMENT). Table 2 lists the product, marketing, supply chain, management, information technology, and human resource strategies characteristic for each stage of globalization.

Differential characteristics of global development.
No international strategy. The initial stage of many firms' global trade is characterized by the lack of an international strategy with

occasional exporting and importing. Such organizations typically focus on their domestic operations and perceive international transactions as a requirement to support domestic business. Specifically, a firm uses a simple export/import strategy to increase the revenues or decrease costs associated with a domestic operation. As Table 2 indicates, lack of an integrated international strategy typically involves a standardized product manufactured in the firm's home country, focused on a limited customer base, with logistics services required for export/import provided by specialized integrated service providers (ISPs). The business context is transaction driven, with common financial statements providing the only level of integration.

Multidomestic strategy. The second international development stage essentially takes the domestic model and replicates it in various locations around the world. As Table 2 indicates, the marketing, operations, and financial strategies mimic the firm's domestic operations while the management and human resource perspectives have a strong home country bias.

While this approach does not require much coordination, it also does not take advantage of potential synergies and economies of scale. This is increasingly a problem when many of the firm's key customers are global and expect a common global offering. From a logistics perspective, the multidomestic strategy does not share the logistics expertise and relationships across regions.

Global strategy. The third stage of international development is characterized by establishment of integrated marketing and supply chain operations in foreign countries. Global operations may include varied combinations of marketing, sales, production, and logistics. Establishment of local facilities and operations serves to increase market awareness and sensitivity. This is often referred to as establishing *local presence*. The local presence strategy uses local production and distribution often supported by a postponement strategy to customize products. Firms engaged in local presence often restrict their operations to a limited number of geographic areas. At the outset of a local presence strategy, foreign operations will often depend on parent company management and personnel, values, procedures, and operations. However, over time, business

Table 2 Differential characteristics of global development.

Development Stages	*Service Focus*	*Marketing Strategy*	*Delivery Strategy*	*Management Strategy*	*Human Resource Development*
No international strategy	Standard product for local market	Single strategy focused on local market	Direct to customer	Single simple financials	Operated by entrepreneur with limited specialization
Multidomestic strategy	Domestic marketing and delivery	Domestic customers	Collaboration	Transaction driven with integrated financials	Management with "home country" focus
Global strategy	Local market customization	Focused specific market areas which may cross international boundaries	Subsidiaries with local presence	Decentralized operations with local profit responsibility	Limited top management with international experience
Transnational strategy	Global branding and integrated operations	Global customers	Worldwide flow of key resources	Centralized planning in global sites	International training and experience

units operating within a foreign market area will have to adopt local business practices.

This adoption typically means developing unique management, marketing, and sales organizations and may include the use of local business systems. As local presence operations expand, the host country philosophy will increasingly emerge; however, the company headquarters' strategic vision remains dominant. Individual country operations are still measured against home country expectations and standards.

Transnational strategy. The transnational strategy contrasts sharply with operations guided by either a multidomestic or global strategy. The original concept of the globally integrated enterprise was popularized in a *Business Week* article two decades ago under the term "The Stateless Corporation" (The Stateless Corporation, 1990). Management of a truly globally integrated enterprise makes market, location, and resource decisions with little or no regard to national boundaries.

Globally integrated enterprises maintain operations and develop a headquarters structure to coordinate across area operations. "Unlike the multinational – which creates mini versions of itself in markets around the world – the globally integrated enterprise locates work, skills, and operations wherever it makes sense, based on access to expertise, on superior economies, and on the presence of open environments and technologies." Thus, the enterprise is globally integrated in the sense that no specific home or parent country dominates policy. IBM offers a good example of an organization that is striving for global integration (Palmisano, 2008). Senior management likely represents a combination of nationalities. Denationalized operations function on the basis of local marketing and sales initiatives and are typically supported by world-class manufacturing and logistics operations. Product sourcing and marketing decisions can be made across a wide range of geographical alternatives. Systems and procedures are designed to meet individual country requirements and are aggregated as necessary to share knowledge and for financial reporting. A truly transnational firm employs global brands with limited customization to reflect market sensitivities, operates in most global regions, employs a global resource view in terms of production and logistics, and incorporates integrated reporting systems and planning technologies to achieve global operating synergies. Although limited customization is desirable to minimize product complexity, the globally integrated enterprise can discriminate between features that represent a critical market difference and those that do not.

Consider, for example, an enterprise that has its historical origin in Germany, Japan, or the United States, but with a high percentage of its sales, ownership, and assets maintained and managed in China. China is estimated to be the world's third largest economy, but there are many supply chain aspects, such as logistics capabilities and infrastructure that are still quite third-world. China's communications, intermodal transport systems, tracking and tracing, and limited highways outside major cities make it very difficult to employ twenty-first century supply chain operating practices. For these reasons, a globally integrated enterprise operating in China needs a combination of local management to facilitate local operations and enterprise management that fully understands the implications of developing business systems, a rapid rate of change, and exploding but unbalanced export/import volume.

Examples of firms that fit the specification of globally integrated enterprises are ABB (Switzerland), Coca-Cola (United States), Dow Chemical (United States), Hoechst (Germany), IBM (United States), ICI (Britain), Johnson & Johnson (United States), Nestlé (Switzerland), Novartis (Switzerland), and Philips (Netherlands). These firms are characterized by a combination of global brands produced and marketed globally with integrated systems and management that can synthesize global operations while being sensitive to regional and local considerations.

While most enterprises engaged in international business are operating in stages one and two, to become a global player a truly international firm must migrate toward global marketing and operations. Such globalization requires a significant level of management trust that transcends countries and cultures. Such trust can

grow only as managers increasingly live and work across cultures.

Managing the global supply chain. To enhance a firm's global capabilities, logistics management must consider five major differences between domestic and international operations: (i) performance cycle structure, (ii) transportation, (iii) operational considerations, (iv) information systems integration, and (v) alliances. These considerations must then be incorporated into the firm's global operating strategy.

Performance cycle structure. The length of the performance cycle is a major difference between domestic and global operations. Instead of 1- to 5-day transit times and 2- to 10-day total performance cycles, global operational cycles often require weeks or months. For example, it is common for automotive parts from Pacific Rim suppliers to take 60 days from order release until physical delivery at a US manufacturing facility. Similarly, fashion merchandise may take anywhere from 30 to 60 days from the time the manufacturer order is released until it is received at a US distribution warehouse.

The reasons leading to a longer order cycle to delivery cycle are communication delays, financing requirements, special packaging requirements, ocean freight scheduling, slow transit times, and customs clearance. Communication may be delayed by time zone and language differences. Financing delays are caused by the requirements for letters of credit and currency translations. Special packaging may be required to protect products from intransit damage due to high humidity, temperature, and weather conditions. Once a product is containerized, it must be scheduled for movement to and between ports having appropriate handling capabilities. This scheduling process can require up to 30 days if the origin and destination ports are not located on high-volume traffic lanes or the ships moving to the desired port lack the necessary equipment. Transit time, once the ship is en route, ranges from 10 to 21 days. Port delays are common as ships wait for others to clear harbor facilities. Customs clearance may further extend the total time. Although it is increasingly common to utilize electronic messaging to preclear shipments through customs prior to arrival at international ports, the elapsed performance cycle time is still lengthy (*see* DIGITAL MEDIUM AND GLOBAL MARKETING). Another problem is restricted availability of containers. Movement from Asia to the United States is generally unbalanced as more material is imported into the United States than is exported to Asia. As a result, there is a strong demand for containers to move product from Asia to the United States but little motivation to ship the empty containers back. This demonstrates how unbalanced trade, either domestically or internationally, can introduce complexity into logistics operations.

These factors cause international logistics performance cycles to be longer, less consistent, and less flexible than in typical domestic operations. This lack of consistency makes planning and coordination more difficult. Determination of shipment status and anticipation of arrival times also requires substantially more effort. The longer performance cycle also results in higher asset commitment because significant inventory is in transit at any point in time.

Transportation. The US initiative to deregulate transportation during the early 1980s has extended globally. The four significant global changes that have occurred are (i) intermodal ownership and operation, (ii) privatization, (iii) cabotage and bilateral agreements, and (iv) infrastructure constraints.

Historically, there have been regulatory restrictions regarding international transportation ownership and operating rights. Transport carriers were limited to operating within a single transportation mode with few, if any, joint pricing and operating agreements. Traditionally, steamship lines could not own or manage integrated land-based operations such as motor or rail carriers. Without joint ownership, operations, and pricing agreements, the operation and tracking of international shipping was very complex. International shipments typically required multiple carriers to perform a single freight movement. Specifically, government rather than market forces determined the extent of services foreign-owned carriers could perform. Although some ownership and operating restrictions remain, marketing and alliance arrangements among countries have

substantially improved transportation flexibility. The removal of multimodal ownership restrictions in the United States and in most other industrialized nations served to facilitate integrated movement. In response to some of these changes in national ownership requirements, an increasing number of global service providers have been established such as DeutschePost, FedEx, TNT, and United Parcel Service.

A second transportation influence on global operations is increased carrier privatization. Historically, many international carriers were owned and operated by national governments in an effort to promote trade and provide national security. Government-owned carriers are typically subsidized and often place surcharges on foreign enterprises that use these services. Artificially high pricing and poor service often made it costly and unreliable to ship via such government-owned carriers. Inefficiencies and inflexibility also resulted from strong unionization and work rules. The combination of high operating cost and low efficiency caused many government-owned carriers to operate at a loss. A great many such carriers have been privatized and must operate in a competitive environment. Carrier privatization has resulted in increased availability of efficient international carriers.

Changes in cabotage and bilateral service agreements are the third transportation influence impacting international trade. Cabotage laws require passengers or goods moving between two domestic ports to utilize only domestic carriers. For example, water shipment from Los Angeles to New York was required to use a US carrier. Similar cabotage laws restricted Canadian drivers from transporting a back haul load to Detroit once a shipment originating in Canada was unloaded in Texas. Cabotage laws were designed to protect domestic transportation industries even though they also served to reduce overall transportation equipment utilization and related efficiency. The European Community has relaxed cabotage restrictions to increase trade efficiency. Such reduced cabotage restrictions will save US corporations 10–15% in intra-European shipping costs. While the US has not rescinded its cabotage laws relating to Canada and Mexico, some of the restrictions have been reduced to enhance equipment utilization and to reduce the environmental impact.

Many regions, both developed and undeveloped, are experiencing major constraints on their physical infrastructure. Global operations are significantly increasing the demand specifically on port and airport capacities. Since much of the infrastructure in the developed world was developed over 50 years ago, it was designed for substantially less capacity and without extensive growth capability in terms of surrounding land. Although information and handling technology has facilitated rapid movement of goods through ports and airports, the volume increase still results in substantial congestion (Norek and Isbel, 2005). At the same time, tight local, state, and federal budgets have limited the reinvestment that is made in the existing infrastructure. As a result, logistics managers are being driven to seek out alternative suppliers, carriers, or port locations.

Operational considerations. There are a number of unique operational considerations to support global supply chains. First, international operations typically require multiple languages for both product and documentation. A technical product such as a computer or a calculator must have local features such as keyboard characters and language on both the product and related manuals. From a logistics perspective, language differences dramatically increase complexity since a product is limited to a specific country once it is language-customized. For example, even though Western Europe is much smaller than the United States in a geographic sense, it requires relatively more inventory to support marketing efforts since separate inventories may be required to accommodate various languages. Although product proliferation due to language requirement has been reduced through multilingual packaging and postponement strategies, such practices are not always acceptable. Some consumers are reluctant to accept products not labeled in their native tongue. In addition to product language implications, international operations may require multilingual documentation for each country through which the shipment passes. Although English is the general language of commerce, some countries require that transportation and customs documentation

Table 3 Common forms of international logistics documentaion.

- *Export irrevocable commercial letter of credit.* A contract between an importer and a bank that transfers liability or paying the exporter from the importer to the (supposedly more creditworthy) importer's bank.
- *Bank draft (or bill of exchange).* A means of payment for an import/export transaction. Two types exist: transaction payable on sight with proper documents (*sight draft*) and transaction payable at some fixed time after acceptance of proper documents (*time draft*). Either type of draft accompanied by instructions and other documents (*but no letter of credit*) is a documentary draft.
- *Bill of lading.* Issued by the shipping company or its agent as evidence of a contract for shipping the merchandise and as a claim to ownership of the goods.
- *Combined transport document.* May replace the bill of lading if goods are shipped by air (*airway bill*) or by more than one mode of transportation.
- *Commercial invoice.* A document written by the exporter to precisely describe the goods and the terms of sale (similar to a shipping invoice used in domestic shipments).
- *Insurance certificate.* Explains what type of coverage is utilized (fire, theft, water), the name of the insurer, and the exporter whose property is being insured.
- *Certificate of origin.* Denotes the country in which the goods were produced to assess tariffs and other government-imposed restrictions on trade.

be provided in the local language. This increases the time and effort for international operations, since complex documents must be translated prior to shipment. These communication and documentation difficulties can be somewhat overcome through standardized electronic transactions.

The second global operational consideration is unique national accommodations such as performance features, technical characteristics, environmental considerations, and safety requirements. Performance feature differences include specific product functionality such as speed or process constraints. Technical characteristics include power supplies, documentation, and metrics. Environmental considerations include chemicals that can be used or the types and amount of waste generated. Safety requirements include automatic shutoffs and specialized documentation. While they may not be substantial, the small differences between country requirements may significantly increase required stock keeping units (SKUs) and subsequent inventory levels.

The third operating consideration is the sheer amount of documentation required for international operations. While domestic operations can generally be completed by using only an invoice and bill of lading, international operations require substantial documentation regarding order contents, transportation, financing, and government control. Table 3 lists and describes common forms of international documentation.

The fourth operating consideration is the high incidence of countertrade and duty drawback found in some international situations. While most established firms prefer cash transactions, countertrade is important. Countertrade, in essence, is when a seller agrees to accept products as payment or purchase products from the buyer as part of a sales agreement. While such agreements have financial consequences, they also have major implications for logistics and marketing in terms of disposal of goods received as payment. Duty drawback describes situations when a firm pays a duty to import goods into a foreign country but the duty paid can be drawn back or returned if the items or a comparable designate is exported. For example, Pepsi supplies syrup to the Soviet government, which bottles and markets the soft drink with practically no control from Pepsi. In return, Pepsi is paid for the syrup by receiving exclusive rights to distribute Russian Stolichnaya vodka in

the United States. This exclusive right requires marketing and logistics support.

Information systems integration. A major challenge in globalization is information systems integration. Since firms typically globalize by acquisition and merger, the integration of systems typically lags. Operational integration requires the ability to route orders and manage inventory requirements electronically throughout the world. Development of supportive technology integration represents substantial capital investment. The overall process was significantly facilitated by the global initiative to achieve Y2K compliance. Two types of system integration are required to support global operations. The first is a global transaction or enterprise resource planning (ERP) system. The global ERP system is necessary to provide common data regarding global customers, suppliers, products, and financials. It is also necessary to provide common and consistent information regarding order and inventory status regardless of the location from which a global customer is inquiring or where the shipment is to be delivered. The second system integration requirement is a global planning system that can maximize overall manufacturing and delivery asset utilization while meeting customer service requirements. Few firms have fully integrated global information systems or capability.

Alliances. A final international operations consideration is the growing importance of third-party alliances. While alliances with carriers and specialized service suppliers are important in domestic operations, they are essential for international commerce. Without alliances, it would be necessary for an enterprise operating internationally to maintain contacts with retailers, wholesalers, manufacturers, suppliers, and service providers throughout the world. International alliances provide market access and expertise and reduce the inherent risk of global operations. The number of alternatives, breadth of activities, and complexity of globalization require alliances.

In summary, globalization is an evolving frontier that increasingly demands more extensive supply chain integration. As firms expand their focus toward international markets, demand for logistical competency increases because of longer supply chains, more variation, increased uncertainty, and more documentation. While the forces of change push toward borderless operations, supply chain logistics management still confronts market, financial, and channel barriers. The barriers are exemplified by distance, demand, diversity, and documentation. The challenge is to position an enterprise to take advantage of the benefits of global marketing and manufacturing by developing world-spanning logistical competency (Cudahy, Narendra and Cases, 2008).

Summary

Global operations is becoming more of the norm for logistics and supply chain executives. Decisions regarding global sourcing and marketing require more complex trade-off analyses than traditionally required for domestic logistics. Both the quantitative and qualitative factors are more complex. While transportation, inventory, and warehousing costs are very substantial for global operations, other cost components, including taxes, tariffs, duties, documentation, and import restrictions, can also have a substantial impact on true total cost. However, in addition to the quantitative considerations, international operations introduce a number of other variables that are much more difficult to quantify. Many of these variables relate directly to logistics operations. The major qualitative considerations include relationship management, infrastructure consistency, production and transit reliability, and security. With increased global marketing and manufacturing operations, logistics management needs to be more involved in developing and implementing global strategies.

Bibliography

This article is adapted from Bowersox, D.J. Closs D.J. and Cooper, M.B. (2010) *Supply Chain Logistics Management*, Chapter 12. McGraw-Hill Publishing, Burr Ridge.

Cudahy, G., Narendra, M. and Cases, C. (2008) Mastering global operations in a multipolar world. *Supply Chain Management Review*, **12** (2), 22–29. The

Council of Supply Chain Management Professionals (CSCMP) has produced a number of *Global Perspectives* reports that describe the logistics characteristics and challenges associated for over twenty specific countries. The reports are available free to members in paper and electronic form from the website at www.cscmp.org. For a more comprehensive discussion, see:

Holstein, W. (1990) The Stateless Corporation. *Business Week*, May 14, 98.

Norek, C. and Isbel, M. (2005) The infrastructure squeeze on global supply chains. *Supply Chain Management Review*, **9** (7), 18–24.

Palmisano, S. (2008) The global consumer gives small companies a big reach. *Financial Times*, May 6, 11.

international advertising – is there still a standardization versus local adaptation debate?

Charles R. Taylor

Introduction

For nearly 50 years, a fundamental issue in the field of international advertising has been the degree to which advertising can be standardized. Even in recent years, standardization versus local adaptation of advertising is often positioned as a debate. A close look at the literature, however, reveals that the idea of a debate in which there is a "standardization camp" and a "local-adaptation camp" is quite dated (Taylor, 2005).

The weight of the research evidence suggests that, at least for large multinational firms, the use of GLOBAL MARKETING STRATEGY, including the creation of a global brand image, has major advantages and can lead to improved performance (Zou and Cavusgil, 2002). One need only to look briefly at the list of the world's best global brands and their value to realize the dramatic advantage a global brand can provide (Interbrand, 2008). For example, in 2008, the top five brands were Coca Cola, IBM, Microsoft, General Electric, and Nokia. The top four of these brands were all valued at over $50 billion. The symbols for these and many other famous global brands are so well entrenched with the average consumer, that they can be flashed on a screen for a period too short to actually read the company name or acronym, yet they are recognized and comprehended by consumers.

In the context of advertising, it has also been shown that given a reasonable level of market similarity the use of a standardized advertising strategy is positively linked to a firm's strategic performance as well as its financial performance. The research suggesting that advertising strategy should be standardized does not in any way minimize the fact that cultural differences still matter. However, it does suggest that multinational advertisers should start the planning process by attempting to set similar broad strategies, including positioning statement and key selling proposition. Clearly, some aspects of execution will likely have to be adapted in many cases, such as the language of the message, specific format of the commercial, actors, celebrities, or spokespeople used (*see* MARKETING CHANNEL STRATEGY).

A good analogy for planning global advertising strategy is to think of the process of manufacturing an automobile. The chassis of a given model can be the same for virtually all consumers. Meanwhile, a few different types of engines may be necessary to please consumers with different desires (performance vs. fuel economy), and numerous options are offered to tailor the product to the consumer's individual preferences. Global advertising can be thought of the same way, in that every attempt should be made to standardize broad aspects of strategy such as positioning of the product (*see* POSITIONING ANALYSIS AND STRATEGIES), while some advertising tasks such as choosing broad media strategy may involve the need for a few modifications, and still others, such as the language of the commercial may need to be modified in many places.

In recent years, scholarly research has shown striking evidence that global advertising strategy pays off. The core reason for this is that frequently cited advantages of standardization pay off.

In examining international advertising from an academic perspective, it is now necessary to take into account both global perspectives on marketing theory at a strategic level as frameworks for understanding the role of culture at a more tactical level. In the following, frequently cited advantages of standardizing advertising versus adapting advertising to local markets are discussed (Czinkota and Ronkainen, 2008).

Advantages of Standardized Advertising

Levitt's (1983) controversial article on the globalization of markets and discussion of the need to standardize the MARKETING MIX as a whole provided important perspectives on some key advantages of standardization. One of the advantages alluded to by Levitt is the achievement of economies of scale. In the context of advertising, this applies most directly to the need to develop and implement advertising campaigns more efficiently. Clearly,

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

for a company operating in, say, 200 markets, developing a completely unique advertising campaign for each of the markets would be inefficient and cost prohibitive. Thus, developing a general strategy and, perhaps, even some tactical aspects of a global advertising campaign centrally can lead to cost savings. Additionally, at least in some contexts, there may be cost savings associated with making media buys that cut across borders, such as by advertising on EuroSport or Sky Channel (*see* MARKETING CHANNEL STRATEGY).

It has become increasingly clear that the establishment of a uniform global brand image is an additional advantage of standardization. Thus, a second key advantage of standardized advertising is that it can be used to help build a uniform brand image around the world. In this way, consumers recognize brands like Starbucks, Samsung, Toyota, Gucci, or Siemens regardless of where they are in the world. With increased evidence that it is possible to target cross-national market segments, it becomes even more important to build a common image while building brand equity cross-nationally (Ko *et al.*, 2007).

It can be argued that a third key benefit of standardized advertising is that it is appropriate for the modern world – described as a global village by Levitt in his discussion. As noted by Levitt, improvements in transportation and communications have led the trend toward consumers having more similar wants and needs than in the past. While Levitt may have exaggerated the degree to which convergence is occurring, there is little doubt that more consumers than ever desire similar types of high-quality goods and services. In such an environment, standardized advertising can capitalize on situations in which wants and needs are similar cross-nationally.

One final benefit of standardization that has been cited is the ability to better coordinate strategy. In the context of advertising, this refers to advertising strategy and campaigns being centrally planned so that they do not end up being in conflict and providing a conflicting image (*see* GLOBAL MARKETING STRATEGY). Simply stated, it is easier to plan and implement a strategy centrally as opposed to doing so in many different counties. At least, in the latter case, more coordination is needed and more oversight is necessary.

Advantages of Local Adaptation of Advertising

Clearly, some aspects of consumer behavior and consumer wants and needs remain influenced by culture. While one can find Dunkin Donuts locations in many parts of the world at a time when more European and US citizens are eating sushi and/or curry than ever before, it is, nevertheless, clear that cultural preferences and tastes play a role in many contexts. Thus, the ability to adjust to specific cultural preferences is important. In an advertising context, this may often involve a need to adapt the tone, format, length, or appeal type used in an advertisement. For example, it is well documented that "hard-sell" advertisements are not as well received in Japan as in some other parts of the world. While it is often possible to retain a single positioning statement to build the brand or general strategic elements of the message (e.g., which feature is being emphasized), quite frequently, there will be elements of the execution that must be altered.

A related advantage of adaptation is the ability to be understood in a local cultural context. This means both communicating in the appropriate language and also in terms understood by citizens of a country or region. Clearly, stubbornly running foreign language advertisements will simply result in failure in many advertising contexts. Perhaps less obvious, however, is the need to choose spokespeople or endorsers the audience in a county can directly relate to. This is easy to understand in the context of the use of celebrity endorsers. Most US citizens exposed to a reel of advertisements in which Korean celebrity endorsers are used to appeal to Koreans would not even be aware that celebrities were appearing in the advertisement.

A third important advantage of adaptation is that it is sometimes necessary to comply with legal restrictions or self-regulatory guidelines in a country. For example, Taylor and Raymond (2000) found that even among countries in the same region (East Asia), regulatory guidelines can vary a great deal, especially for certain types of products. For example, in Japan it is permissible to advertise the alcohol content

of a product and even refer to it as "*extra strong*," a term that would not be allowed in many other countries. Alcoholic spirits may be advertised in some countries but not others. Similarly, DTC prescription drug advertising is banned in all OECD countries except the US and New Zealand, but other forms of promotion are allowed in various markets. When regulations vary across countries, advertisers may be obliged to adapt their strategy.

Key Empirical Studies of Standardization versus Adaptation

While a large volume of studies have been conducted on standardization and related issues, the findings of these studies have tended to converge. In a comprehensive review of the literature, Taylor and Johnson (2002) conclude that many companies have adopted an approach where broad strategies are standardized, but advertising executions are localized as needed. In a key study in the area using survey methodology, Harris (1994) found that most multinational companies use advertising, but the extent to which they use it varies. He emphasizes that the practitioners should focus more on "how" to standardize as opposed to "whether" to standardize.

Subsequent studies appear to verify Harris' findings. In a major survey of multinational advertisers, Duncan and Ramaprasad (1995) found that strategy was most often standardized by multinational advertisers, while it was used less often in execution and least often in language. Notably, the authors found that the desire to achieve a uniform brand image was the number-one motivator for standardization – higher than the motivation provided by potential cost savings.

In an examination of whether advertising standardization by multinational corporations pays off in the form of improved financial and strategic performance, Okazaki, Taylor, and Zou (2006) conducted a survey of Japanese and US multinational firms operating subsidiaries in the European Union. Results showed a positive link between standardization and both financial and strategic performance. Again, however, the study provided evidence that strategic elements are more likely to be standardized than executional elements.

While there has been some convergence in the research evidence as to how standardization is and should be used by multinational corporations, the findings suggest that theories pertaining to both global marketing as well as conceptual frameworks help understand and measure culture. If standardization is typically important at a strategic level, it is important to apply theoretical or conceptual perspectives that allow us to better understand how and why standardization strategies work. However, it is also clear that understanding and implementing standardization alone does not allow for fully effective global advertising. As a result, cultural and country-specific factors must be taken into account.

Fortunately, some recent advances in theoretical perspectives have allowed for better application to the issue of degree and effectiveness of standardized advertising.

Global Marketing Theories

At least three major perspectives have begun to emerge that are aimed at better understanding the global market and how GLOBAL MARKETING STRATEGY work. A first theory is the global marketing strategy theory put forward by Zou and Cavusgil (2002). Drawn from industrial organization theory as well as additional perspectives, this theory proposes a model that outlines eight aspects of a global marketing strategy. The eight dimensions are product standardization, promotion standardization, distribution standardization, pricing standardization, coordination of value-adding activities (integration of competitive moves), global market participation, coordination of marketing activities, and concentration of marketing activities. Global marketing strategy theory argues that the degree of fit, referring to the degree to which a company's own global marketing strategy matches up with the environment the firm faces, as well as it strategic capabilities, will have a profound impact on success.

Global marketing strategy can be applied to the context of advertising because it is a key element of promotion in many instances, and

because it also reinforces product/branding strategy. Moreover, it clearly must be coordinated with other marketing-mix activities. Notably, in their study, Zou and Cavusgil (2002) found a positive relationship between following a global marketing strategy and financial and strategic performance under circumstances where the fit was sufficient.

A second influential theoretical perspective in the global marketing realm is global consumer culture theory (Arnould and Thompson, 2005). This theory argues that a trend toward a global consumer culture has indeed taken place and that market segments that cut across national boundaries can often be identified. In such instances, it should be possible for the multinational firm to improve performance. However, it is also notable that the global consumer culture theory avoids simply positing that consumers behave similarly around the world. On the contrary, the theory argues that similarities and differences can exist both within a culture and across cultures. This theory has also been applied to international advertising, and can be used to help explain some of the nuances associated with designing effective international advertising.

A third relevant perspective is an outgrowth of global consumer culture theory and can be referred to as the *global brand-positioning theory*. This work is an outgrowth of global consumer culture theory and was begun by Alden, Steenkamp, and Batra (1999). The basic idea is that the growth of global consumer culture has led to the development of a transnational consumer culture. In this environment, multinational firms realize an enhanced ability to target across countries the segments and consumers who have shared consumption values (Terpstra and David, 1991). A key finding from this line of research is that perceived brand globality is positively related to consumers' perceptions of the quality and prestige of the brand (*see* BRAND VALUE). Moreover, firms operating successfully in more regions benefit from higher global reach, which further contributes to a competitive advantage. In sum, this theory suggests that companies can achieve success by developing uniform positioning across markets and by achieving synergy as a result of operating across multiple markets (*see* POSITIONING ANALYSIS AND STRATEGIES).

In the context of advertising, global brand-positioning theory argues for the need to advertize a brand in multiple markets in a way that reinforces a single image. In this way, the benefits of higher perceptions of quality and prestige can be realized by raising perceived brand globality.

Cultural Dimensions

The above theories have contributed to our understanding of how global strategies can be used successfully in advertising. Yet, as mentioned, there is a clear need to at least take cultural and country-specific measures into account in planning advertising executions. At least three major perspectives have had influence in the academic literature and they also have practical applicability.

The first perspective is that of the anthropologist Edward T. Hall, who contributed the idea of high-context versus low-context cultures to the literature. The difference in communication styles captured by Hall has been examined as it applied to advertising in several studies, many of which have shown that low-context cultures (e.g., the United States) tend to use more direct styles of communication in advertising, whereas less direct styles are often preferred in high-context cultures (e.g., Japan). Hall's classification of societies in terms of time orientation, in which some cultures have a monochromic time orientation while other have a polychromic orientation, has also been influential, though it has been applied less frequently to advertising.

A second and highly influential perspective on culture that has been widely applied to advertising research is Geert Hofstede's dimensions of culture (Hofstede, 1980). Indeed, Hofstede's original four dimensions of culture: individualism/collectivism, uncertainty avoidance, power distance, and masculinity/femininity have been widely studied in terms of their influence on differences in advertising. While some general criticisms of the Hofstede dimensions and the method by which they were derived have been made, there is no question that this perspective has contributed some insight into the subtle differences in advertising in different countries. For example, several studies have shown that there are differences in the types of

advertising appeals that work in individualistic versus collectivistic cultures. Power distance has been used to examine the depiction of rank and formality in advertising in various cultures. Studies have also applied uncertainty avoidance in terms of attitudes toward advertising of certain types of products (e.g., pharmaceuticals and over-the-counter drugs) and appeals. While masculinity/femininity has been applied less frequently, it has been included as a dimension in several content-analysis studies. Hofstede's long-term/short-term dimension, which was added later, has also been tested in some studies.

Recently, House and his colleagues (2004) have developed a competing set of cultural dimensions that has begun to be applied to advertising research. On the basis of a large-scale study of 62 countries, House identified nine cultural dimensions: performance orientation, in-group collectivism, institutional collectivism, power distance, uncertainty avoidance, future orientation, humane orientation, assertiveness, and gender egalitarianism. While this scheme and its methodology are also not without detractors, it is also beginning to be applied to advertising studies. From an advertising perspective, there is significant overlap with Hofstede's dimensions. However, dimensions of assertiveness and performance orientation have been of particular interest to researchers. It appears that the future will see more applications of this framework.

The three perspectives listed above are by no means the only cultural frameworks that have been applied to advertising. Several others, notable among them being the one suggested by Harry Triandis, have contributed important perspectives that advertising researchers have picked up on. As culture is a highly complex construct, it seems likely that work in studying and better understanding cultural dimensions and their impact on advertising will continue for many years. There have been many calls for a more general theory of culture as it applies to culture, but to date such a theory has been quite elusive, perhaps confirming to just how complex culture really is.

Summary and Conclusion

While standardization versus local adaptation is no longer categorized as topic of debate, there is some merit in both perspectives. Research has demonstrated that global strategies can have advantages and, in most instances, marketers and advertisers are well advised to pursue global branding strategies reinforced by global advertising. At the same time, as one moves from strategies to executions, cultural differences and other country-specific issues do indeed matter, so localization issues must be considered. The weight of the evidence argues for considering global strategies first, realizing that, at the same time, executional modifications will often need to be made.

Future research will contribute more theoretical insight into global advertising. It is likely that this will occur both at a strategic level, with a focus on global strategies, and a more tactical level with a focus on cultural differences.

Bibliography

Alden, D.L., Steenkamp, J.-B.E.M., and Batra, R. (1999) Brand positioning through advertising in Asia, North America, and Europe: the role of global consumer culture. *Journal of Marketing*, **63** (1), 75–87.

Arnould, E.J. and Thompson, C.J. (2005) Consumer culture theory (CCT): twenty years of research. *Journal of Consumer Research*, **31** (4), 868–892.

Czinkota, M. and Ronkainen, I. (2008) *International Marketing*, 8th edn, Thomson/Southwetern, Cincinatti.

Duncan, T. and Ramaprasad, J. (1995) Standardized multinational advertising: the influencing factors. *Journal of Advertising*, **24** (3), 55–68.

Hall, E.T. (1976) *Beyond Culture*, Anchor Press/Doubleday, Garden City.

Harris, G. (1994) International advertising standardization: what do the multinationals actually standardize? *Journal of International Marketing*, **2** (4), 13–30.

Hofstede, G. (1980) *Culture's Consequences: International Differences in Work-related Values*, Sage Publications, Beverly Hills.

House, R.J., Hanges, P.J., Javidan, M. *et al.* (2004) *Culture, Leadership and Organizations: The Globe Study of 62 Societies*, Sage Publications, Beverly Hills.

Interbrand (2008) Best Global Brands: 2008 Rankings, www. interbrand. com/ best_ global_ brands. aspx (accessed February 8, 2009).

Ko, E.J., Kim, E., Taylor, C.R. *et al.* (2007) Cross-national market segmentation in the fashion industry: a study of European, Korean, and US consumers. *International Marketing Review*, **24** (5), 629–651.

Levitt, T. (1983) The globalization of markets. *Harvard Business Review*, **61** (3), 92–102.

Okazaki, S., Taylor, C.R., and Zou, S. (2006) Advertising standardization's positive impact on the bottom line: a model of when and how standardization improves financial and strategic performance. *Journal of Advertising*, **35** (Fall), 17–33.

Taylor, C.R. (2005) Moving international advertising research forward: a new research agenda. *Journal of Advertising*, **34** (Winter), 7–16.

Taylor, C.R. and Johnson, C.M. (2002) Research on standardization vs. specialization of advertising in the 1990s: a review and directions for further research. *Advances in International Marketing*, **12**, 45–66.

Taylor, C.R. and Raymond, M.A. (2000) An analysis of product category restrictions in advertising in four East Asian markets. *International Marketing Review*, **17** (3), 287–304.

Terpstra, V. and David, K. (1991) *The Cultural Environment of International Business*, 3rd edn, Southwestern, Cincinatti.

Zou, S. and Cavusgil, S.T. (2002) The GMS: a broad conceptualization of global marketing strategy and its effect on firm performance. *Journal of Marketing*, **66** (4), 40–56.

global marketing strategy: perspectives and approaches

Susan P. Douglas and C. Samuel Craig

Introduction

Markets worldwide are becoming increasingly integrated across national borders at a macroeconomic, competitive, and product market level. Firms of all sizes, and in almost all industries, are increasingly conceptualizing their strategy on a global basis. As a result, increasing interest in studying this trend and in understanding its key characteristics has developed. As noted by Zou and Cavusgil (2002) a number of different approaches have been adopted, which vary in terms of their theoretical or conceptual underpinnings, and focus on different facets of marketing strategy as well as their definition of terms such as *global* and *marketing strategy* (*see* GLOBAL MARKETING STRATEGY). This has resulted in the absence of a generally accepted conceptualization of global marketing strategy, and hence, an ability to generalize findings from different research studies and more broadly improve understanding with regard to the impact of globalization on the firm's competitive position.

In this article, a framework is presented that aims to provide a clear understanding of global marketing strategy and at the same time allows incorporation of other perspectives. First, the different terms that are used in the present article are explained. Next, the dominant approaches to studying international/global marketing and the theoretical perspectives that underlie these approaches are briefly reviewed and the key topics covered. The various issues confronting the firm as it develops a global marketing strategy are then discussed, starting with the initial development of the strategy, followed by the ongoing process of refining and developing that strategy and consolidating the strategy to improve efficiency and competitiveness in global markets.

Key concepts.

Global. The term *global* is used to define the geographic scope of the firm's operations and strategy development. A firm is considered to have a global marketing strategy if it is involved in marketing its products and services in most geographic regions and areas in the world, has established a clear strategy as to how these operations are managed in each of these areas, and retains control over how these operations are managed and evaluated (*see* GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES). This does not necessarily imply that these operations are globally integrated or standardized on a worldwide basis. The term *global* will be used broadly to refer to any involvement outside the firm's home market, as global strategy must start somewhere.

Marketing strategy. A *marketing strategy* is defined as a strategy that is based not only on identifying target customer needs and interests in a clearly defined product market in order to develop customer value creation but also in clearly identifying the firm's distinctive skills and capabilities relative to those of other competitors in the marketplace (*see* COMPETITIVE ADVANTAGE: ITS SOURCES AND THE SEARCH FOR VALUE). This results in the development of a marketing strategy based on the firm's competitive advantage and skills in relation not only to marketing activities such as new product development, pricing, advertising, and distribution, but also to other elements of the value chain, both upstream and downstream such as design, production, sourcing, and logistics.

Alternative Approaches and Perspectives to Global Marketing Strategy

In studying global marketing strategy, a number of different approaches have been adopted ranging from the transaction cost perspective to the evolutionary and global integration perspective. Typically, each focuses on different decisions or aspects of global marketing strategy and corresponds in many respects to differences in the experience of the firm in international markets.

The transaction cost approach. One of the earliest approaches adopted in studying the development of global marketing strategy was transaction cost analysis (Anderson and Gatignon, 1986). This focused on the

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

appropriate choice of mode of entry into international markets and viewed such decisions as a trade-off between control and the cost of resource commitments, and was grounded in Williamson's (1981) transaction cost economics. While control enables the firm to coordinate actions, execute and revise strategies, and thus obtain a higher return, it also entails commitment of resources and hence exposure to risk in an uncertain environment. This perspective has subsequently been widely used in assessing mode of entry decisions (Erramilli and Rao, 1983), notably for exporters and service firms.

The standardization/adaptation perspective. The standardization/adaptation issue (*see* STANDARDIZATION/ADAPTATION OF INTERNATIONAL MARKETING STRATEGY) was initially raised by Buzzell (1968), examining the potential benefits of standardizing different elements of the marketing mix as opposed to adopting localized strategies. This became a central debate characterizing the marketing literature following Levitt's controversial article "The Globalization of Markets" Levitt (1983), which argued that multinational firms would only be successful if they marketed standardized products worldwide, taking advantage of potential economies of scale in production, distribution, marketing, and management. This debate has been widely pursued not only in relation to marketing strategy in general (Douglas and Wind, 1987) but also in relation to the benefits and feasibility of standardization relative to different elements of the marketing mix, products, and target segments (Jain, 1989).

The global configuration/coordination perspective. This perspective emphasizes the importance of configuring and coordinating the firm's activities at different stages in the value chain across different countries so as to improve efficiency and gain the maximum competitive advantage (Craig and Douglas, 2000; Takeuchi and Porter, 1986; Roth, 1992). Activities at the upper end of the value chain such as sourcing, design, and engineering should, for example, be concentrated in countries where they can be performed most effectively and cost efficiently. At the same time, activities should be both vertically and horizontally coordinated at different stages in the value chain and across countries to optimize cost efficiency and maximize speed of response to changes in demand or competitor moves.

The global integration perspective. Another perspective is the global integration approach (Yip, 1995; Zou and Cavusgil, 2002). World markets are viewed as an integrated whole, and emphasis is placed on the importance of conducting operations in all major markets worldwide, and integrating strategy development and execution across these markets (*see* FORCES AFFECTING GLOBAL INTEGRATION AND GLOBAL MARKETING). Similarly, resources may be shifted from one market to another in order to compete more effectively. For example, competitive attacks in one market may be met by counterattacks in a competitor's home market or other key markets. In addition, emphasis is placed on developing a strategy for the standardization of product, promotion, and distribution activities across world markets.

The evolutionary perspective. The most comprehensive perspective is to view the firm's operations as evolving over time as the firm gains experience and expands in international markets. The "stages" theory of internationalization developed by Johansson and Vahlne (1977), based on a study of the pattern of internationalization of Swedish firms, argues that the perceived risk associated with international expansion leads firms to enter proximate, more familiar markets first, gradually expanding into more distant market as experience is gained in operating in international markets. Similarly, Douglas and Craig (1989, 1995) suggest that international market expansion can be viewed as a sequential decision-making process starting with decisions relating to entry into international markets, standardization versus adaptation of international marketing mix decisions to different environmental conditions, and subsequently focusing on coordination and integration of these decisions across national markets.

More recently, it has been pointed out that some firms are "born global" (Knight and Cavusgil, 1996, *see* BORN GLOBAL). In essence,

such firms immediately adopt a global perspective in initial market entry, and target customers worldwide. Typically, these firms target "niche" markets where customers in different countries have similar needs and interests, for example, medical equipment and computer software. The growth of the international communications infrastructure, particularly, the Internet, often facilitates identification of these opportunities and establishment of relations with customers, ensuring that customer needs are met and satisfied on an ongoing basis.

Since the evolutionary framework also provides the broadest perspective and can also incorporate the other perspectives, this article also focuses on the specific decisions that a firm needs to make as it develops experience in international markets.

Beginning Global Market Operations

While not all firms expand operations to the point where they can be considered global, all begin by entering a country or countries outside their home market. In initially entering into global markets, a firm needs to make three key decisions:

- which countries to enter;
- what modes of operation to adopt; and
- the timing and sequencing of entry.

These decisions need to be considered in the light of the firm's objectives with regard to global markets, particularly with regard to the desired degree of involvement, and amount of resources (human and financial) a firm is willing to devote to developing operations in international markets, as well as the level of risk – macroeconomic, competitive, and policy – that it is willing to consider in going international.

Selecting countries to enter. In deciding which countries to enter (*see* MARKET ENTRY AND EXPANSION), the firm needs to first evaluate opportunities on a worldwide basis, assessing in each country the macroeconomic environmental factors, such as population size and growth, the level of GNI (Gross National Income) and rate of economic growth, the degree of urbanization, the rate of inflation, the level of corruption, political risk, financial risk, trade barriers, and market regulation. In addition, opportunities and risk at the product market level need to be assessed. Here, the firm needs to consider the absolute size of the product market and its rate of growth, as well as per capita consumption and growth. If a product is not currently marketed in a country, surrogate indicators of demand need to be used. The level of competition also needs to be considered as also the presence and strength of other global, regional, and local competitors. Markets that are large may appear attractive, but if the rate of growth is low this may signal that the market is saturated. Similarly, the market may have a high rate of growth, but if there is a substantial degree of competition, the firm may prefer to focus on developing operations and stimulating primary demand in a less competitive market. In conducting this analysis, a hierarchical screening process based on secondary data can be used to assess opportunities worldwide in order to reduce and expedite the assessment of opportunities and risk.

Mode of operation. Once having assessed opportunities and determined which countries to enter or consider entering, the firm needs to consider the mode of operation, that is, whether to enter via exporting (*see* EXPORT PERFORMANCE), a contractual agreement with another firm such as licensing or franchising, contract manufacturing, or joint venture, or alternatively, to set up a wholly owned subsidiary either on a greenfield basis or via acquisition. Here, a key factor is the degree of control a firm wishes to exercise over operations in international markets as well as the importance of local input and experience in managing and developing operations, and the resources a firm is willing to devote to international expansion. Some firms, for example, wine or agricultural producers, have no choice but to export if they wish to enter international markets. Similarly, service operations, for example, fast food, hotel chains, and car hire typically expand via franchising to avoid the cost of acquiring local facilities and to ensure input of local management capabilities and experience (*see* INTERNATIONAL FRANCHISING). This, however, typically requires extensive systems of control and training on a global or regional basis to ensure that franchisees in all markets provide

a consistent service experience and reinforce the firm's brand image internationally.

Timing and sequencing of global market entry. The firm also needs to decide the *timing* of its entry into different international markets, that is, whether to be a first mover and enter a country ahead of competition, or alternatively be a fast follower (*see* INTERNATIONAL PRODUCT DIFFUSION). Being a first mover enables the firm to establish its brands and develop customer and distributor loyalty ahead of competition as well as to monopolize key resources such as strategic locations, key distribution channels, thus erecting entry barriers for competition. On the other hand, being a first mover also entails substantial risks as the firm needs to expend resources to stimulate primary demand and in some cases, develop the market infrastructure besides convincing distributors to stock the product. Similarly, there is substantial uncertainty as to whether potential demand will develop or whether environmental conditions will change, for example, the imposition of product regulations. Similarly, competitors can learn from the firm's mistakes and "leapfrog" the learning curve, entering later with a more desirable product or effective marketing strategy.

Another aspect is the *sequencing* of market entry. A firm may decide to enter international markets sequentially, adopting, for example, a "rehearsal" strategy, entering one country in a region first to gain experience in that market, then rolling out into other countries. For example, a US firm entering the European market might decide to enter the Netherlands first, and use that experience to develop strategies for entering other large markets in Europe, for example, France, Germany, and the United Kingdom. A variant of this strategy is to enter a single country in each region first and use this as a base for developing operations and strategy in other markets in the region. For example, McDonalds first entered into the United Kingdom, and then expanded into the rest of Europe, and used Australia as a basis for developing operations in Southeast Asia.

A key factor driving strategy in initial entry is the desire to achieve *economies of scale*. Typically, a firm will use the same marketing strategy, particularly in terms of products, product positioning, and branding in international markets as in domestic markets so as to achieve *tangible and intangible* economies of scale. *Tangible* economies typically arise from production economies of scale or spreading R&D and other investment costs, for example, in developing advertising themes and copy, over a larger volume of sales. *Intangible* economies may be less apparent, arising from use of the brand or corporate image on an international basis, thus enhancing its visibility and value to customers. Other intangible economies may include specialized management skills and know-how, for example, in the management of franchise operations or the development of creative product and marketing ideas as in Apple's product design capabilities, or Bic's skill in developing disposable products.

While many firms, both large and small, have already entered international markets, typical firms that are still in the early stages of entering international markets are

- *entrepreneurial* firms (*see* INTERNATIONAL ENTREPRENEURSHIP), often targeting niche markets worldwide, for example, firms selling specialized medical equipment, export services, or new product variants such as soft serve ice cream;
- *large emerging market multinationals* that often enter markets in developed countries with a low cost positioning based on resource cost in their home market, for example, Haier, a Chinese consumer electronics firm; or Embraer, a Brazilian manufacturer of small jet aircraft; or Mahindra and Mahindra, an Indian manufacturer of agricultural equipment.

Refining and Developing Global Marketing Strategy

Typically, once firms have entered a number of countries, they begin to expand within these markets. This is stimulated in part by concern with meeting local competition and meeting specific local needs and interests. This is further reinforced by local management attitudes and initiatives that typically reflect a belief that

local market characteristics and demand conditions are different and require adaptation of products and marketing strategy. An important concern is also more effective utilization of local knowledge and assets resulting in product line extension or development of new products that can make use of existing channels of distribution. Constraints imposed by national market barriers, for example, tariff barriers, quotas, import duties, and local product regulation, may also encourage local production.

In entering and expanding within countries that the firm has decided to enter, the following three decisions are of paramount importance:

- how far to market *standardized products worldwide* or adapt to differing conditions in different countries (*see* STANDARDIZATION/ADAPTATION OF INTERNATIONAL MARKETING STRATEGY), and similarly, whether to adopt the same marketing strategy, that is, pricing, promotion, and distribution strategy or to adapt locally;
- whether to *extend the product line* to include local variants adapted to local customer demand and market conditions;
- whether to *develop new products* to meet specific local needs, for example, in emerging markets or markets with different climatic conditions (*see* INTERNATIONAL PRICING OBJECTIVES AND STRATEGIES; DESIGNING A GLOBAL SUPPLY CHAIN: OPPORTUNITIES AND CHALLENGES).

In considering whether or not to standardize or adapt different elements of the marketing mix, the firm needs to consider a number of different factors. While there are a number of benefits from standardization, there are also barriers to standardization and advantages to adapting, and these need to be weighed relative to each element of the marketing mix. Some mix elements such as product and how it is positioned may be more readily standardized than other elements such as pricing (*see* INTERNATIONAL PRICING OBJECTIVES AND STRATEGIES) or distribution (*see* DESIGNING A GLOBAL SUPPLY CHAIN: OPPORTUNITIES AND CHALLENGES). Similarly, the degree of modification may vary ranging from minor modifications such as adjusting to voltage, size or color preferences, to major differences such as taste and technological complexity.

One of the primary *benefits of standardization* is that firms are able to achieve economies of scale in R&D and production. Product standardization in international markets enables a firm to reduce the number of models worldwide and hence reduce development costs in addition to enabling them to afford a higher level of design expertise. Similarly, pharmaceutical companies are able to spread the high cost of developing new drugs, and consumer goods companies the costs of advertising (*see* INTERNATIONAL ADVERTISING – IS THERE STILL A STANDARDIZATION VERSUS LOCAL ADAPTATION DEBATE?) development across a higher sales volume. Standardization also enables firms to transfer ideas, experience, and knowledge developed in one market to other markets, for example, the concept of disposable products, or experience in developing environmentally friendly products and packaging (*see* GLOBAL CONSUMERISM AND CONSUMPTION). Standardized products also facilitate the development of a uniform image of quality and service and the ability to develop a strong corporate or brand image worldwide. Standardization also facilitates coordination and control of operations in different country markets as uniform performance standards can be established and performance more readily compared across countries.

On the other hand, there are a number of *barriers to standardization*. Differences in customer characteristics and response patterns caused by different sociocultural values and lifestyles, climatic or usage conditions, or perceptions and associations with different images, may generate need for different products, or promotional and distribution strategies. Similarly, government regulations and restrictions relating particularly to product and promotional decisions, or campaigns to buy local products, may result in the need to adapt product positioning and promotional strategy. Likewise, differences in the marketing infrastructure, for example, the availability, cost, and effectiveness of different media such as TV, radio, and print as well as nontraditional media, such as the Internet or viral marketing have to be accounted for. Similarly, the organization of the distribution infrastructure and the importance of large-scale

distribution organizations versus small Mom and Pop stores may vary considerably from one country to another, or the presence of international or regional retailers may vary considerably, facilitating or hindering local brands and products. Similarly, the extent of and strength of local or regional competitors may vary. Typically, the presence of strong local competition will create pressures to adapt either in terms or product, promotion, or pricing strategy. In addition, local managers are typically opposed to standardized products as these reduce their control over marketing strategy, as marketing policy and guidelines are likely to be established at regional or corporate headquarters.

Such differences in market and demand conditions in different countries and regions, typically lead the firm to *extend the product line* adding new product variants that are adapted to specific needs and tastes (*see* MANAGING THE GLOBAL PRODUCT PORTFOLIO). This may include adding product versions with new flavors or scents, or different types of soft drinks, or bottled waters. This typically helps to generate additional sales, fills out the firm's product line, and enables the firm to tap a broader range of customers as well as compete more effectively in the local marketplace.

At the same time, the firm may also *develop new products* tailored to specific local market needs (*see* INTERNATIONAL PRODUCT INNOVATION AND DEVELOPMENT). For example, in emerging markets the firm may develop simple functional products, for example, mobile phones targeted to low-income consumers who are unable to afford more complex high-end products marketed in developed countries. Similarly, computers that uses icons may be designed for illiterate consumers, and solar-powered equipment created for consumers with no or unreliable access to electricity. Again, this will help the firm's sales growth in a given market and tap a broader base of customers.

In extending the product line and developing new products within a given country, a key priority is to leverage *economies of scope*, adding products and product variants that can be distributed through existing channels of distribution or produced at existing production facilities. This enables the firm to spread overhead costs over a higher volume of products. At the same time, it utilizes experience and knowledge of market conditions in a given country, and investment in research to identify and understand customer needs. Similarly, where the same channels of distribution are used, relations developed with distributors or sources of supply may be further reinforced. In essence therefore, a key aim is *building* the firm's operations in a given market, and particularly in building the scale and scope of operations so as to compete more effectively in local markets.

Firms focusing on refining and developing their international marketing strategy, are typically firms competing predominantly on a regional basis, for example, companies such as Henkel, the German-based manufacturer of detergents and other household products, as well as products targeted to handymen such as glue, paint, and solvents. Similarly, Kao, the Japanese detergent and personal products manufacturer, is focusing on growth in Southeast Asian markets. Other companies are transitioning from developing their strategy to consolidating their positioning and strategy across regional and global markets. McDonald's, for example, has rolled out local variants such as shrimp burgers and fried rice patties developed in Japan to other countries in Asia and McArabia flat bread in Europe. Similarly, in Europe many facilities are being upgraded so that customers come not only for fast service and inexpensive food, but also to enjoy a comfortable and clean environment. Other service ideas, such as home delivery started in Cairo, Egypt, are being added in other busy city centers.

Consolidating/Integrating Global Strategy

A number of factors act as triggers to consolidate and rationalize operations across different markets (*see* FORCES AFFECTING GLOBAL INTEGRATION AND GLOBAL MARKETING). These include, for example, cost inefficiencies and duplication of effort across different country markets, particularly those that are similar in terms of demand characteristics or are geographically proximate. Similarly, opportunities may occur for the transfer of products and brands developed in one country market to other country markets, targeting similar market

segments, for example, high-income consumers, younger consumers, in business-to-business markets, or global customers. Similarly, the emergence of competition on a global scale is facilitated by improved linkages between national market infrastructures, for example, retailing or advertising media.

Such factors lead firms to pay increased attention to consolidating operations in global markets, and improving the coordination and integration of operations at different levels of the value chain, such as promotional efforts through greater use of regional or global media and positioning strategies or the establishment of global and regional product development centers. Similarly, sourcing or production strategies may be coordinated or configured at a global level. Here, however it is important to note that global configuration of sourcing and production strategies are dependent on standardization at the product design level.

A key element of the firm's strategy at this point is therefore to establish a *global portfolio of countries, products, and target segments* in order to establish direction for future efforts (*see* MANAGING THE GLOBAL PRODUCT PORTFOLIO). Then, it needs to establish global strategy based on the market scope and target market of its various product business, which may be focused on a specific target segment worldwide, or alternatively be broad based, targeting the mass market in different countries, or hybrid, that is, some combination of both. Finally, the firm will need to consider improved integration and coordination of operations both upstream and downstream in the value-chain as well as horizontally across geographic regions and product businesses.

In establishing the geographic scope of the global portfolio, the firm should maintain a balance between mature markets, such as the United States, Western Europe, and Japan, while targeting emerging growth markets such as China, India, Brazil, Thailand, or Colombia. The mature markets provide low growth, but also lower risk and require little additional investment to develop and build demand. The emerging markets, on the other hand, offer high growth potential, but require greater investment to understand customer demand, evaluate the nature of local competition, and build the market infrastructure, particularly in underdeveloped regions. At the same time, the degree of integration across markets, for example, in Europe, South America, or the Indian subcontinent needs to be considered, in order to effectively allocate investment efforts within a region, as well as diversify across different regions of the world.

Once the firm has established the countries and regions to target, the next step is to establish the firm's GLOBAL MARKETING STRATEGY. Here, an important issue is how far this is integrated across different countries and regions of the world. This is likely to depend, to a substantial degree, on the scope of the product market. Where the firms adopt a focused strategy targeting a specific market segment, such as high-income consumers, or young adults, they are likely to integrate strategy across markets, adopting the same positioning, the same or similar product line and pricing, promotional copy and media, and distribution strategy. This may result in the use of global or regional media and, similarly, global or regional retailers who are able to provide coverage matching the geographic scope of the firm's operations and hence, improve marketing efficiencies. If, on the other hand, the firm targets a broad-based mass market, considerably greater difficulties may be encountered in integrating strategy across markets. While increasingly, many firms are adopting global branding strategies, the extent of the product line may vary from one region to another or even by country. Similarly, while advertising themes may be uniform across countries and regions, their execution may vary and distribution strategies may similarly need to be adapted to differences in the structure of the distribution system and the availability and reach of organized distribution, for example, supermarkets, mass-market retailers, and department stores.

In integrating and consolidating strategy across countries and regions, a primary concern, as in initially entering international markets, is to achieve both *tangible and intangible scale efficiencies*, particularly with regard to the management of marketing and service operations. At the same time, *synergies* may arise from coordinating and integrating strategy and operations, especially across proximate markets, or in integrating communications

efforts at a regional or global level. Similarly, the *transfer of best practices* is critical, that is, ideas for products, promotional or distribution strategies across countries and regions, as well as experience and know-how in effectively managing marketing operations in different environmental conditions.

Many large US and European multinationals are in the process of consolidating operations across world markets. Many consumer goods firms, who previously had local or regional brands are focusing on building a global branding strategy, at the corporate, product business, and product level. In some cases, this leads to the development of tiered branding strategies where the product level brand is endorsed by the corporate brand as well as a family or house brand. Attention is also paid to building a global information system, particularly at the firm level so that production and distribution logistics can be better coordinated and integrated across countries and regions as well as across product businesses. Increasing emphasis is also paid to the transfer of management across geographic regions in order to enable them to develop experience in working in different environmental contexts, and hence develop and train a multicultural global workforce of managers capable of operating in a variety of different contexts. At the same time, they are able to bring their experience working in other countries and regions to deal with similar problems and situations in a given country or context.

Issues in the Continuing Evolution of Global Marketing Strategy

The global marketing strategy of firms continues to evolve in response to a changing environment and new challenges. Three of the most pressing issues are (i) the *increasing complexity* of managing operations on a global scale; (ii) coping with the *increasing intensity of competition* not only from established multinationals but also from new competitors from emerging markets; and (iii) *exploiting opportunities in emerging markets*. The role of each of these in shaping the evolution of global marketing strategy is further elaborated.

Complexity of managing international operations. As international operations grow in importance and complexity, management has to direct, coordinate, and control operations on a much broader and more diverse scale and scope (*see* MARKETING STRATEGY IMPLEMENTATION). This may entail decisions relating to the reconfiguration of the geographic organization of operations at different stages in the value chain, for example, developing global production platforms, or centers of product innovation, or improving vertical and horizontal coordination at different stages of the value chain as well as developing external communication linkages with customers, developing global account management systems, improved supplier linkages, and organizational restructuring to improve coordination and communication links across countries and regions. In addition, establishment of a global information system, and of a global management workforce, consisting of managers able to operate in different countries and cultural contexts are essential to operate effectively in increasingly culturally diverse and far-flung world markets.

Increasing intensity of competition. Firms have also to cope with the increasing intensity of competition as well as the emergence of new sources of competition. As growth slows in many markets in developed countries, such as the United States, Western Europe, and Japan, competition between established multinationals in these markets has become increasingly severe. This has been exacerbated by the entry of firms from emerging markets such as China, India, or Brazil. These firms are able not only to leverage the advantages of a low-cost resource base to build their position in global markets and enter developed markets but are also learning to adopt the technologies and management skills of firms in developed countries and in some cases surpass them in terms of innovative skills. Typically, these firms enter developed markets with a low-price positioning. However, often, as they establish a market position and effective distribution channels, they begin to move upward, adding higher-priced products and developing their brand image. This poses an increasing challenge to established multinationals in developed countries.

Exploiting opportunities in emerging markets. Slower growth in developed markets has prompted firms to look to emerging markets for growth opportunities. Although current income levels in these countries are low relative to developed countries, ranging from $4460 in Russia to $730 in India, the growth rates and economic fundamentals in these countries suggest immense future market potential (*see* EMERGING MARKETS). Both India and China have sizable and relatively affluent middle classes which constitute a large market for goods and service as well as vast markets of rural poor and an increasing number of highly successful entrepreneurs. Here, an important decision for developed market firms is whether to focus on the affluent urban middle class in these countries, leveraging existing products and global brands, with limited adaptation of the marketing mix, or alternatively, target lower income consumers, typically in rural areas. Targeting lower income consumers, however, typically requires the development of radically new marketing strategies, including the development of new low-cost functional products and improvement of distribution access and efficiency in rural areas. In addition, creative ways to enhance the ability of consumers to afford products need to be developed. While requiring substantial effort and ingenuity, the size and potential of the emerging markets, not only in India and China but also other continents such as South America and Africa, offers tremendous opportunities for future growth.

CONCLUSION

The crafting of global marketing strategy is a dynamic ongoing process, continually evolving as the firm expands into new countries and markets, requiring adaptation to new market conditions and demand factors, competitive forces, as well as internal pressures within the firm. Also, the fundamental nature of global marketing strategy changes as the firm's involvement in global markets increases. To succeed, the firm must become an organism that continually evolves, adapts, and responds to the changing realities of the global marketplace. Firms that are able to do so will prosper; those that do not will wither.

Bibliography

Anderson, E. and Gatignon, H. (1986) Modes of foreign entry; a transaction cost analysis and propositions. *Journal of International Business Studies*, **1**, 1–26.

Buzzell, R. (1968) Can you standardize international marketing?. *Harvard Business Review*, **69**, 102–113.

Craig, C.S. and Douglas, S.P. (2000) Configural advantage in global markets. *Journal of International Marketing*, **8** (1), 6–25.

Douglas, S.P. and Craig, C.S. (1989) Evolution of global marketing strategy: scale, scope and synergy. *Columbia Journal of World Business*, **24** (3), 47–58.

Douglas, S.P. and Craig, C.S. (1995) *Global Marketing Strategy*, McGraw Hill, New York.

Douglas, S.P. and Wind, Y. (1987) The myth of globalization. *Columbia Journal of World Business*, **22** (4), 19–29.

Erramilli, M.K. and Rao, C.P. (1983) Service firm's international entry-mode choice: a modified transaction cost analysis approach. *Journal of Global Marketing*, **57**, 19–38.

Jain, S. (1989) Standardization of international marketing strategy: some hypotheses. *Journal of Marketing*, **53**, 70–79.

Johansson, J. and Vahlne, J.-E. (1977) The internationalization process of the firm- a model of knowledge development and increasing foreign market commitments. *Journal of International Business Studies*, **8**, 47–58.

Knight, G. and Cavusgil, S.T. (1996) The born global firm: a challenge to the traditional theory of internationalization. *Advances in International Marketing*, **8**, 11–26.

Levitt, T. (1983) The globalization of markets. *Harvard Business Review*, **61**, 92–102.

Roth, K. (1992) International configuration and coordination archetypes for medium-sized firms in global industries. *Journal of International Business Studies*, **23** (3), 533–549.

Takeuchi, H. and Porter, M.E. (1986) The strategic role of international marketing: managing the nature and extent of worldwide coordination, in *Competition in Global Industries* (ed. M.E. Porter), Harvard Graduate School of Business Administration, Cambridge.

Williamson, O.E. (1981) The economics of organizations: the transaction cost approach. *American Journal of Sociology*, **87** (3), 548–577.

Yip, G.S. (1995) *Total Global Strategy: Managing for Worldwide Competitive Advantage*, Prentice Hall, Englewood Cliffs

Zou, S. and Cavusgil, S.T. (2002) The GMS: a broad conceptualization of global marketing strategy and its effect on firm performance. *Journal of Marketing*, **66**, 40–56.

digital medium and global marketing

Venkatesh Shankar

Introduction

The digital revolution is transforming the business landscape and profoundly influencing marketing in an increasingly global environment. From a demand perspective, the digital medium or the Internet (e.g., World Wide Web, email) has offered firms access to new customers, markets, and business models across the globe. From a supply standpoint, the digital medium has enabled firms to cut costs of marketing and operations by coordinating their value chains around the globe. In this article, we focus on the role of the digital medium or the Internet on the global marketplace and global marketing.

The role of digital medium or the Internet in global marketing decisions and the impact of the Internet on firm performance in global markets can be analyzed through an organizing framework (Shankar and Meyer, 2009). This framework addresses the following important questions. How does the Internet affect a firm's global marketing decisions? What are the direct and indirect effects of the Internet and Internet marketing strategy on firm performance in global markets?

Companies use the Internet in the global marketing context in different ways. Firms can use the Internet for gathering information, providing customer support, and improving customer relationships. Some firms use the Internet as a primary information source and information dissemination vehicle to perform global market research and to identify customer segments that span different countries. Others use it as a medium for communicating a brand's value proposition or position to its target audience across countries. Broadly speaking, companies use the Internet to formulate and implement global marketing mix decisions.

Shankar and Meyer's (2009) organizing framework relating to the Internet, global marketing decisions and firm performance is shown in Figure1. The global/international marketing mix decisions include those on product, brand, price, communication, promotion, and distribution channels. The Internet and Internet marketing strategy directly influence both the global marketing mix decisions and the firm's performance. The Internet and the Internet marketing strategy of a firm also have moderating effects on the impacts of each global marketing mix decision on firm performance. Because firm performance is critical to firms, we focus on the direct and moderating effects of the Internet and Internet marketing strategy on firm performance.

In formulating their global digital marketing strategy, firms can compare different countries on dimensions such as infrastructure, geographical distance, language, buyer behavior, buyer demographics, country image, payment systems, and currency using a framework based on two dimensions: global integration and local responsiveness (Guillen, 2002). Depending on the combination of these dimensions, he recommends four global Internet marketing strategies: pure local adaptation, global cost leadership, nationally differentiated, and transnational cost adaptation strategies. According to him, each strategy is appropriate for specific product categories. According to him, for example, products whose features are most amenable to direct inspection, such as clothing, cars, and collectibles should follow a nationally differentiated strategy because these categories need high local responsiveness in website design, language, return policy, and customer service, but low integration across countries. The framework is useful for classifying different categories but offers few guidelines on leveraging the digital medium in the global marketing context.

Digital Medium and Global Product Development

The Internet is increasingly used in global product development. Companies can use real time collaboration software for product design so that product developers across the globe can connect and simultaneously work on the same product idea. A driving factor for using the Internet in global product development is shorter design cycles fueled by the opportunity to develop products on a 24 × 7 basis globally. Companies use global Web-based design platforms to develop products through collaborative teams across the world. The primary benefits

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

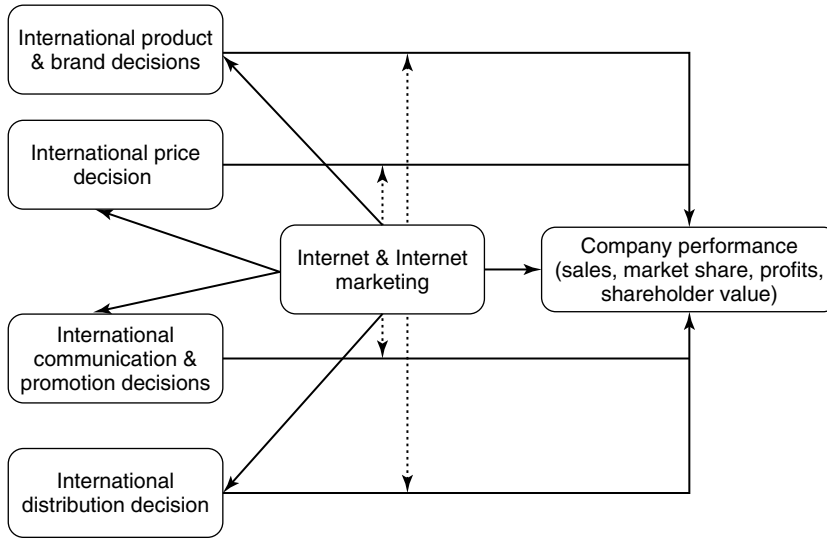


Figure 1 An organizing framework for Internet and international marketing. (Bold lines represent direct effects. Dashed lines represent moderating effects.) Source: Shankar and Meyer (2009).

of Web-based global product development are reduced product development time, greater ideas and inputs from design engineers around the world, and better time-leveraging of talent located at different time zones. Some companies also use these Web-based systems to work across the globe with "offshoring" partners.

The Internet also plays an important part in the diffusion of a new product within and across countries. For products such as pharmaceutical drugs and movies, the Internet serves as a powerful medium to inform potential users and customers across countries. On the one hand, firms can leverage this information dissemination ability of the Internet to accelerate the diffusion of their products across countries. On the other hand, if customers in the initial markets had adverse experiences with their products, firms may be hampered by the Internet in new global markets. Therefore, firms need to more carefully plan the design and management of product launches in the initial markets.

To sum up, the Internet has an important role to play in product development across countries. Managed appropriately, the Internet can result in better new product ideas, more effective collaboration, shorter development cycle time, and better use of talent across multiple time zones. The Internet can play both positive and negative roles in the diffusion of new products across cultures and countries. Hence firms will have to more carefully plan their product introductions in the initial countries.

Digital medium and global communication. The Internet plays a key role in companies' communication and promotion efforts and in their effectiveness in the global marketplace. Communication efforts can be of two types: company-generated and user-generated. Company-generated communication efforts are typically centered on company websites. User-generated communication efforts relate to activities such as the creation and management of social media, community sites, blogging, and file sharing by customers.

Company-generated communication. Culture affects customer attitudes toward company websites, and hence has a strong effect on website effectiveness. Most companies have country-specific websites. For example, Procter & Gamble created regional sites during the 2006 World Cup soccer championship to promote its Gillette, Braun, Duracell, and Oral B brands

and to raise awareness of its status as an official sponsor (Shankar and Meyer, 2009). Users from several countries first selected one of the four geographical regions on its website and then had the option to choose the language in which the website appeared. Research suggests that local language and local adaptation are keys to the success of global marketing on the Internet.

An important strategic issue related to the Internet in the global context is the globalization versus localization of products and websites. Globalization refers to the standardization of products and sites across countries and cultures, while localization refers to the adaptation of products and sites to different countries (Shankar and Donato 2003). While companies should naturally adapt their websites to local languages, the extent to which they should adapt the website content to the individual countries would depend on the product development and marketing costs, culture-specificity of products, importance of brand equity, and the degree of country-specific customer needs.

User-generated communication. The Internet enables users to share information and create global communities focused on specific topics. User-generated communication in the global context can be classified into different forms: posting on social media such as Facebook, MySpace, Twitter, and LinkedIn; blogging on own as well as community global sites; podcasting; posting videos on video sites such as YouTube and Flickr; and posting product reviews in global communities. By measuring and monitoring such communication about the firm and its products, a firm can use the information to better manage its relationship with its customers and improve its products and customer service. For example, before the launch of PlayStation 2, a global brand community that allowed consumers to discuss and anticipate attributes of the new product had emerged. However, because the Internet also allowed "brand terrorists" (users who can control a brand in ways detrimental to the firm owning the brand), Sony decided to launch its own global brand community so that it could monitor and proactively listen to the conversations among consumers. Another example is Stormhoek winery in South Africa. Through the use of various online marketing activities, including blogging, Stormhoek increased its shipments to the United Kingdom from 50 000 cases in 2005 to 350 000 cases in 2007 (Business Day, 2007).

Another example of a global online brand community is NikePlus, designed with music collaboration from Apple, that offers an array of useful tools for running enthusiasts. These tools include managing own runs, issuing running challenge to friends, socializing with other community members across the world, obtaining music through Apple, and sharing information through blogs.

To summarize, the Internet moderates the effect of communication and promotion on firm performance in the global context. By better understanding customer needs across different countries and cultures, firms can develop appropriate content on their websites in different markets. They could also measure and monitor user-generated communication and proactively use the information for improving products, enhancing customer service, and deepening customer relationships.

Digital Medium and Global Pricing

The Internet affects prices and their dispersion across sellers. The Internet has allowed different segments to become aware of prices offered to one another, regardless of where the segments are physically located. For example, the pricing of pharmaceuticals in Europe is changing such that price differentials across countries are narrower because consumers know that the price of a drug in Spain is different from that in Belgium. Sometimes, online price dispersion across countries may be influenced by regulatory authorities. Consider the case of pricing of Apple's *iTunes* in Europe (Sweeny, 2008). Until 2008, the prices for downloading a song or album through *iTunes* were higher in the United Kingdom than in 16 other European countries. For example, in France and Germany, music buyers were charged €0.99 (74p) per track, while British music fans were charged 79p. Following consumer complaints, the European Commission investigated Apple for unfair pricing. In early 2008, Apple agreed to reduce the price it charges UK users to buy

tracks from *iTunes* by almost 10% within six months to bring them in line with the rest of Europe. Although Apple finally made the decision to follow a harmonized pan-European pricing policy, its ability to do so also depends on the willingness of the major record labels to adopt a pan-European standardized view of pricing.

These examples highlight the role of price transparency in firms' pricing strategy across global markets. Although the Internet has brought increased transparency on costs and prices, it has also allowed firms to highlight and differentiate on nonprice attributes (Shankar, Rangaswamy and Pusateri, 2001). It is possible for firms to tailor their offerings to the needs of consumers in different countries or offer branded variants across countries, thus reducing the inclination or ability of customers to directly compare prices of the same item across countries.

Owing to such possibilities, there are differences among prices and dispersion of prices among retailers across different countries. Ancarani *et al.* (2008) argue that on the one hand retailer price levels and dispersion may be similar across countries because channel competition and the roles of channels are increasingly similar across countries and the borderless nature and transparency of the Internet can have a positive influence on the similarity of retailer pricing across countries. However, on the other hand, they suggest that retailer price levels and dispersion may be different across countries because of differences in the adoption rate of the Internet, consumer attitudes toward the Internet, price sensitivities, and competitive landscape across countries.

Ancarani *et al.* (2008) present an empirical analysis of retailer price levels and dispersion using data collected for different product categories (e.g., books, CDs) in three European countries, namely, France, Germany, and Italy. Their results show that, in general, price levels, including shipping costs, are higher online than offline in each of these three countries and that price dispersion is persistent across these countries. Multichannel retailers have the highest price levels in each of these countries, but they do not exhibit the highest price dispersion. Their results suggest that the opportunities for price differentiation for a given type of retailer may be different in different countries. Their data, however, are restricted to two product categories in three Group 7 (G7) countries and may not be generalizable across developing economies.

Thus, consumer and company use of the digital medium have important influences on firm prices and on the effect of pricing on firm performance. The Internet enhances price transparency and allows customers to compare prices across countries. However, empirical analysis of price levels and price dispersion suggests that price dispersion is persistent, and the opportunities for price differentiation do exist and may be different across countries.

Digital Medium and Global Channels

The Internet serves as a distribution channel for many firms for several products. It often acts as a direct distribution channel for marketers of items ranging from apparel to computer hardware and software to books to CDs and DVDs to electronic equipment. In some cases, the Internet serves as a substitute channel for other distribution channels such as physical stores and catalogs. In other cases, it acts as a complementary channel. The use of the Internet as an important distribution channel in the emerging practice of multichannel marketing is growing.

In the global context, the use of the Internet as a distribution channel is significant because it allows many firms to reach a wide global audience without substantially increasing the cost of channel development. However, the practice of multichannel marketing in the global context depends on the degree of substitute or complementary effects of the Internet relative to other channels in each country. In countries where the complementarity of the Internet with other channels is high, firms will practice greater multichannel marketing than in countries where the Internet is perceived as a substitute to other channels.

A firm's extent of use of the Internet as a distribution channel in each country may depend on the country, customer, company, and competition factors. The country factors include regulatory issues, taxes, transportation modes, geographical proximity of the country to fulfillment center, Internet penetration level, and logistical infrastructure. Customer factors

include desired delivery speed, willingness to pay, the extent of physical inspection desired, and the influence of consumer-generated digital media. Company factors comprise market reach goals, distribution competency, fulfillment capability, shipping costs, and the like. Competition factors include the number and intensity of competitors in that country, the distribution channels of competitors in the country, channel expertise of competitors, and anticipated channel moves of competitors (Shankar, Rangaswamy and Pusateri, 2010a, 2010b). Depending on the combination of these factors, a firm may tailor the extent of the use of the Internet as a distribution channel for different countries.

The evidence for the use and success of the Internet as a distribution channel in the global context, however, is mixed. While many firms use their websites as store fronts to customers in multiple countries and fulfill orders that they receive through their sites, because of the level of investment required by the clients, physical market presence and personal contact may be more important for sales. However, information designed for and placed on the Internet can improve a firm's reputation and credibility, making personal selling easier in global markets.

The example of Stormhoek wines in South Africa illustrates how Internet can help global distribution for some types of products. By leveraging UK bloggers to sell directly to UK consumers, Stormhoek became "the wine of the blogging world." Stormhoek's shipments to the United Kingdom increased from 50 000 cases in 2005 to 350 000 in 2007 (Business Day, 2007). A well-designed channel strategy involving the Web across global markets will likely improve firm performance. However, apart from anecdotal evidence, there is sparse research on the effects of the Web as a distribution channel on firm performance across global markets.

In summary, there is mixed evidence on the use of the Internet as a distribution channel in global markets. The use of the Web as a channel depends on factors relating to country, customer, company, and competition. Although the use of the Web as a channel is likely to have a positive effect on firm performance in global markets, there is not enough evidence on this topic for us to make a strong conclusion.

Future Opportunities and Conclusion

As Internet penetration in different countries continues to grow, the role of the Internet in global marketing will keep rising. The impact will be more significant and often more dramatic in countries where Internet penetration is still low and has enormous potential for improvement. In some countries, the ability of the Internet as a viable new medium of communication and channel of distribution can significantly impact economic growth.

A major development related to the Internet is the spread and rise of mobile media and technology across the world. Mobile devices such as cell phones, personal digital assistants, digital music players (e.g., iPod), and hybrid devices (e.g., iPhone, iPad) now provide more pervasive connectivity to websites and users through mobile Internet than ever before. Many developing countries are leapfrogging others in the use of the mobile Internet and email (through short-messaging service or SMS). For example, two emerging economic superpowers, China and India, are major beneficiaries of the surge of mobile Internet. China has the biggest user base of mobile phone subscribers, while India has the fastest growing mobile subscriber base (Shankar and Balasubramanian, 2009). Such rapid penetration of mobile Internet and connectivity will accelerate the impact of Internet marketing activities on firm performance across the world.

The rise in importance of the Internet and the mobile media in the global context offers several opportunities for future research on global marketing issues. Important questions in this regard are how does customer behavior with regard to the use of the Internet vary across countries? How do customers differ in mobile media usage across countries? How does the mobile Internet affect firm's marketing mix decisions? What impact does mobile Internet have on firm performance? What is the impact of user-generated communication among customers across diverse cultures on the diffusion of products across countries?

With regard to measures of firm performance, research on the Internet and global marketing has at best focused on company sales. Future research should examine measures such as profits and shareholder value. The availability of data on

Internet marketing activities in the global context will continue to be a challenge. In particular, because company data on costs and profits by country are confidential, it would be difficult to collect such data. Nevertheless, more empirical research in these areas will offer deeper insights into Internet and global marketing.

Not much is known on the differences between goods and services with regard to the role of the Internet in global marketing. Are the effects of the Internet on global marketing mix decisions and on the relationships between these decisions and firm performance the same for goods and services? In particular, are there differences between digitizable goods and digitizable services? Digitizable products (e.g., books, music, video, software) are those that can be easily distributed over the Internet to customers. In the global context, these products assume significance as they can be downloaded by customers in multiple countries any time. The *iTunes* is an example of such a digitizable product. With the launch and high initial sales of e-readers such as Amazon's Kindle, Barnes and Noble's Nook, and Apple's iPad, which offer advanced reading benefits, how should firms approach global marketing of print content? Future research could address these interesting questions and topics.

In conclusion, the explosive growth in the use of the digital medium continues to alter global marketplace and global marketing in important ways. The digital medium and global Internet marketing strategy have both a direct effect and moderating effects on the impact of marketing mix decisions on firm performance. With regard to global product development, the Internet has significant influences on the effectiveness and speed of new product development and its impact on firm performance. The Internet also has an important role in the effects of both company- and user-generated communication efforts on firm performance. On the global pricing dimension, the Web allows more pricing transparency, but also permits opportunities for differentiation across countries. With regard to global distribution, the Web may serve as either a substitute or a complementary channel in different global markets and by coordinating the Internet with other channels, firms can improve performance in global markets.

In the future, continued Internet penetration and the surging growth of mobile media may change global marketing further. Research on the digital medium and global marketing is still growing and many important questions remain largely underexplored. More research is needed to better understand the relationships among the Internet, mobile Internet, marketing mix decisions, and firm performance in the global context.

See also *competitor analysis; competitive analysis; marketing strategy; marketing strategy models*

Bibliography

Ancarani, F., Frank, J., Frederic, J. and Shankar, V. (2008) *Are Price Levels and Price Dispersion Among Retailer Types Similar Across Countries? A Cross-Country Empirical Analysis*, SDA Bocconi, Italy. Working Paper.

Business Day (2007) Blogging, MXit Challenge Traditional Marketing June 25, 5.

Guillen, M.F. (2002) What is the best global strategy for the internet? *Business Horizons*, May-June, 39–46.

Shankar, V. and Balasubramanian, S. (2009) Mobile marketing: synthesis and prognosis. *Journal of Interactive Marketing*, **23** (2), 118–129.

Shankar, V. and Donato, M.P. (2003) Personalization of global sales and marketing activities in the digital economy, in *Power of One* (eds N. Pal and A. Rangaswamy), eBRC press, Penn State University, University Park, PA.

Shankar, V. and Meyer, J. (2009) Internet and international marketing, in *Handbook of International Marketing* (eds M. Kotabe and C. Helsen), Sage, pp. 451–467.

Shankar, V., Rangaswamy, A. and Pusateri, M. (2001) *The Online Medium and Customer Price Sensitivity*, Penn State University, University Park, PA. Working Paper.

Shankar, V., Rangaswamy, A. and Pusateri, M. (2010a) Competitive analysis, in *Encyclopedia in Marketing*, John Wiley & Sons.

Shankar, V., Rangaswamy, A. and Pusateri, M. (2010b) Competitor analysis, in *Encyclopedia in Marketing*, John Wiley & Sons.

Sweeny, M. (2008) Apple to cut UK prices for iTunes tracks. *The Guardian*, January 8.

market entry and expansion

Keith D. Brouthers

Introduction

Although business expansion into global markets has grown rapidly in the past few decades, research in the area has lagged. Foreign market entry and expansion entails a number of important marketing strategy decisions including (i) selection of specific target countries, (ii) structuring of foreign subsidiary units, and (iii) management of foreign operations. In this article, we focus on the decision about structuring of international subsidiary units, often referred to as *entry-mode choice*. Research in this area has tended to focus almost exclusively on firm- and country-specific characteristics, showing little interest in the effects of marketing strategy, and managerial decision-making behavior (*see* INTERNATIONAL MARKETING CHANNELS). Brouthers and Hennart (2007) provide a detailed review of the entry-mode literature and suggest that scholars have tended to concentrate on rational models of strategic choice, ignoring behavioral aspects. Further, Canabal and White (2008) in their review note that most international entry-mode research has been published in international business and management journals with only a few contributions in marketing research outlets. This may explain why critically important issues involving marketing strategy have been virtually ignored by researchers so far.

International entry-mode choice, like other marketing strategy decisions, is influenced by a multitude of factors. These can be classified into three distinct groups (Figure 1): firm/country characteristics, top management team effects, and marketing strategy fit. The main theoretical perspectives that have been applied to the entry-mode decision – transaction cost analysis, resource-based view (RBV), and institutional theory – focus on firm and country characteristics. These theories do an excellent job of explaining how differences in country characteristics, such as institutional environments (Brouthers, 2002), and firm characteristics, like resource advantages (Brouthers, Brouthers, and Werner, 2008a), lead to specific entry-mode choices. What these theories and related research studies do not do is to help explain how marketing strategy issues as well as managerial decision-making processes and influences play a part in the entry-mode decision.

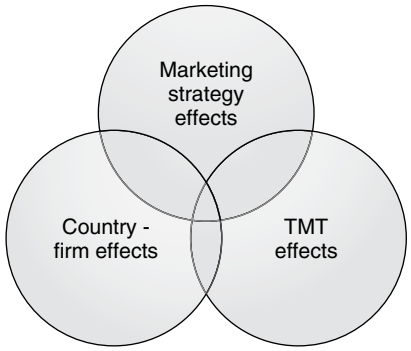


Figure 1 Dimensions of entry-mode choice.

In this article, we provide a short overview of the existing entry-mode research and elaborate on areas where future marketing research can make an important contribution. Before we begin the review, we define some of the key terms and concepts that lie behind this research area.

Key Terms and Concepts

Entry modes – also referred to as *channel choice*, *channel selection*, *modes of entry*, or *entry structures*– vary along a continuum of risk and control (Brouthers and Hennart, 2007). At the high end are wholly owned subsidiary units where a firm owns 95% or more of the equity of the entity, and has made a large financial commitment, which exposes the firm to high risks, but also provides the firm with total control over the foreign operation. At the other extreme, are nonequity license agreements, where firms sell the right to use their product/process to a foreign-based firm, and exporting modes (*see* STRATEGIC EXPORT MARKETING–ACHIEVING SUCCESS IN A HARSH ENVIRONMENT). At this end of the spectrum, the firm has made only a small financial commitment and is therefore exposed to very low risk, but it also has little control over the foreign operation. Between these extremes are multiple variations including equity joint ventures (including majority, equal, and minority ownership), nonequity cooperative agreements (also commonly called

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

strategic alliances), and franchising agreements (a complex form of licensing) (*see* INTERNATIONAL FRANCHISING). As Canabal and White (2008) note, most studies exploring the entry-mode decision look at the choice between two or more of these mode types.

While exporting is an important entry mode, the marketing literature has tended to examine exporting channel selection as a decision separate from entry-mode choice. Scholars like Bello and Lohtia (1995) and others (*see* STRATEGIC EXPORT MARKETING–ACHIEVING SUCCESS IN A HARSH ENVIRONMENT) tend to use the same theoretical approaches in export channel choice as entry-mode scholars use in entry-mode choice. This tends to create some confusion among researchers because the exporting literature tends to be the only market entry and expansion literature where researchers separate the mode choice for the sales function of a firm from the mode choice of the production function of the firm. In fact, most entry-mode studies do not specify whether they are examining the entry for production, sales, or both functions of the firm. Hence, although the terminology is different (entry mode versus channel choice) the types of modes examined and the theories used in these two streams of research tend to coincide.

Existing Research

Firm and country characteristics. Most of the literature in entry-mode choice has been published since 1990. The growth in entry-mode (channel choice) scholarship since 1990 has provided researchers and managers with a better understanding of how firms choose between various entry-mode types. This research tends to use theories that focus attention on firm- or country-specific characteristics. For example, in Canabal and White's (2008) review they note that of the 10 most widely used independent variables, 6 are firm-specific (including international experience and firm size), and 4 are country-specific (including cultural distance and host-country risk).

Three main theoretical perspectives drive the study of entry-mode choice: transaction cost economics (TCE), institutional theory, and the RBV (Brouthers and Hennart, 2007). TCE is the most widely applied theoretical perspective and captures both firm- and country-specific influences. TCE suggests that managers have limited cognitive abilities and hence are concerned with safeguarding their investments against issues arising from incomplete contracts and opportunism. This concern is amplified when the investment involves specific assets (a firm-specific attribute) as well as internal uncertainties (a firm-specific attribute) and external uncertainties (a country-specific attribute). These three characteristics of the transaction then influence the mode-choice decision.

Two recent meta-analyses of TCE entry-mode studies have found that asset specificity, internal uncertainties (also referred to as *behavioral uncertainties*), and external uncertainties (sometime referred to as *host-country risk*) all have a significant impact on entry-mode choice; providing strong support for the transaction cost perspective (Geyskens, Steenkamp, and Kumar, 2006; Zhao, Luo, and Suh, 2004). Further, these studies tend to show that foreign subsidiary performance is better when the mode used is predicted by transaction cost theory. This suggests that TCE can provide a good foundation for managers to develop a tool for determining the structure of new foreign subsidiary units that will result in greater success.

Furthermore, although TCE entry-mode choice models have been extensively tested on both service-based and product-based firms, normally, studies examine only product- or service-based firms, or include a dummy variable to "control" for this industry type. But these studies offer little insight into how decision-making and mode choice may differ between service providers that expand abroad and firms expanding to sell tangible products. There has been some progress in this area using TCE, but much more work is needed.

Probably, the most influential research study comparing service and manufacturing firm mode choice is the one by Erramilli and Rao (1993). They suggest that because services and manufactured products differ, transaction cost decision models of entry-mode choice need to be modified to take these differences into account. They go on to provide a theoretical extension to transaction

cost theory and test the revised theory empirically. Yet recently, Brouthers and Brouthers (2003) looked at the same issue and suggested that no adjustment to transaction cost theory was needed. They maintain that service- and product-based firms simply are influenced to a greater (lesser) extent by the basic transaction cost factors. They extend this work by suggesting the real difference between service- and product-based firms is associated with how risk and trust propensity influence mode choice and transaction-cost- based decisions.

In addition to transaction cost research, other studies have used an institutional theory perspective focusing on country-specific factors that may influence the mode-choice decision (Brouthers and Hennart, 2007). Institutional theory suggests that each country is composed of a unique combination of normative, cognitive, and regulative factors that impact the way business is carried out and how managers, employees, and customers behave. One stream of research in this area suggests that the institutional environment produces an isomorphic effect where firms simply replicate existing entry-mode structures to conform to industry or country standards. A second stream in this area suggests that the individual dimensions of the institutional environment (or the distance between home country and target country dimensions) have differing effects on mode choice; either creating opportunities for entry or restricting access to resources, customers, or distribution channels.

Originally conceptualized as national cultural distance and country risk, entry-mode studies tend to provide growing guidance on how the institutional environment impacts mode choice. The only meta-analysis in this area explores national cultural distance research (Tihanyi, Griffith, and Russell, 2005). This study finds that cultural distance does not tend to have a direct effect on mode choice but that it only has an impact on entry-mode choice for firms coming from high risk propensity countries (like the USA and the UK). This suggests that the cultural component of the institutional environment may have different effects for firms from different home countries entering different host countries. Other aspects of the institutional environment such as host-country risk and legal restrictions have also been examined, but as yet no meta-analysis exists (see Brouthers and Hennart, 2007 for a summary of these studies). In general, research using institutional theory has provided managers with an understanding of how home- and host-country institutional factors influence the choice of entry mode and the success of foreign subsidiary units.

The third most widely used theoretical perspective, the RBV, focuses on firm-specific factors and generally suggests that firms are a collection of resources and capabilities (knowledge and processes) that can be combined in different ways to create a competitive advantage. This advantage can be used to help the firm enter foreign markets and overcome the liability of foreignness. Compared to both transaction cost and institutional-theory-based research, far fewer studies have examined entry-mode choice using the RBV because of the difficulty in defining and measuring resources, many of which, like tacit knowledge, are intangible (Brouthers and Hennart, 2007).

One of the more frequently explored resources that has been examined is experience. Scholars like Anderson and Coughlan (1987) and Delios and Henisz (2000) have noted that different types of experience may provide different benefits to the firm that can have an important impact on entry-mode choice (either in combination with transaction cost factors as noted by Anderson and Coughlan (1987) or moderated by the institutional environment as noted by Delios and Henisz (2000)). Some studies taking an RBV perspective tend to look at how foreign entry, and hence mode choice, can be used to exploit existing resources in new foreign markets while other research examines how the mode choices of firms can be used to help them scout for new resources to enhance the firm's existing stock. Research using this perspective has highlighted the importance of linking the type of resource advantage and motive for foreign expansion (exploitation of existing resource-based advantages or exploration for new resource-based advantages) with an appropriate subsidiary structure (Brouthers and Hennart, 2007).

In addition to studies using a single theoretical perspective, scholars often develop studies that combine multiple theories. For example, Brouthers, Brouthers, and Werner (2008a)

combine the RBV and institutional theory to examine their joint impact on mode choice. They find that resource-based advantages do not always have the same value in foreign markets as in the home market. Therefore, firms may need to vary their mode choice to get the most value out of existing resource-based advantages when entering foreign markets. Brouthers (2002) among others (Brouthers and Hennart (2007)) explore the mutual effects of TCE and institutional theory. They find that institutional theory helps extend TCE by clarifying additional external uncertainties that firms may face when entering new foreign markets. As Brouthers (2002) indicates, by considering both transaction cost characteristics as well as institutional factors, firms may make more informed mode-choice decisions that result in better performing foreign subsidiary units. Many other studies have combined aspects of two or more theoretical perspectives to examine the international entry-mode choices of firms. The one thing these studies all have in common is that they concentrate on theories that focus on firm- and country-specific aspects of the mode-choice decision (see Brouthers and Hennart (2007) for more details).

In sum, despite the growth in studies examining mode (channel) choice, most have used theoretical perspectives that focus on country- and firm-specific aspects of the decision. Although very helpful to managers, this research is limited because it means that we still have little understanding of how issues like marketing strategy or managerial biases influence the mode-choice decision.

Potential for Future Research

There are now almost 200 studies providing theoretical and/or empirical examination of the issue of firm expansion and mode choice. Recently researchers have been questioning the need for further research in this area. Yet despite the growth in research there is still much to be learned. We now discuss two areas of research that have the potential to make a significant contribution to improving international performance through a better understanding of the method of entry-mode choice.

Marketing strategy. Marketing strategy deals with a wide range of decisions that can influence the performance of the firm (*see* GLOBAL MARKETING STRATEGY; GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES). Yet, when examining international expansion, researchers tend to ignore the marketing strategy implications of mode choice. Some research does address strategy issues. For example, Bradley and Gannon (2000) look at how market concentration versus diversification strategies impact mode choice while Sanchez-Peinado, Pla-Barber, and Hebert (2007) examine the impact of global versus multidomestic strategy, follow-the-leader strategies, and market-seeking strategies on mode choice. In these and related papers, researchers have found mixed support; strategies appear to provide an important impact on mode choice but the way that impact is felt is not entirely clear (Brouthers and Hennart, 2007).

Other marketing strategy issues have not found their way into entry-mode research despite their importance to the firm. For example, variations in product/service mix may have an influence on mode choice (*see* MARKETING MIX; MANAGING THE GLOBAL PRODUCT PORTFOLIO). Specific entry-mode types may enhance (detract from) the marketing effort in new foreign markets simply because of the product/service mix being offered. Normally firms do not use the same product/service mix in each foreign market as they use at home. These differences may call for different skills and knowledge, which may impact the mode-choice decision because different modes of entry provide firms with varying degrees of access to skills and knowledge.

Further, the marketing strategy of using standardized or adapted products/services in foreign markets may also influence mode choice (*see* STANDARDIZATION/ADAPTATION OF INTERNATIONAL MARKETING STRATEGY). Standardization, for example, may require few country-specific adjustments to the production process but extensive knowledge and skills in marketing the product/service to the final customer, hence influencing the mode choice for the production operation differently from the sales function. Likewise, adaptation of products/services may require extensive

country-specific production knowledge but might build on existing marketing skills. These differences may have a significant influence on mode choice and the effectiveness of the foreign operation.

Market orientation is now recognized as critically important in domestic markets, yet there is little understanding about the best way to extend this orientation to foreign markets (*see* MARKET ORIENTATION). A firm whose competitive advantage relies on its market orientation may need to think carefully about how to be successful in foreign markets. For example, is it more effective to use wholly owned subsidiaries and keep control of foreign operations when market orientation is important or does sharing ownership through joint ventures or license agreements help firms more effectively service foreign customers?

Branding strategy is also not considered in the mode-choice decision, yet brands are often one of the most powerful marketing tools a firm may possess (*see* BRAND STRATEGY; GLOBAL BRANDING: THREE KEYS FOR GLOBAL BRAND SUCCESS). Brands and brand strategy may significantly impact mode choice because brands may be better protected through some modes than through others. In addition, some mode types may make it easier for firms to obtain additional brands or expand brands to related products/services. Hence brand strategy may be another important determinant of the mode-choice decision.

As we have briefly outlined above, marketing strategy issues form an important and integral part of a firms success, yet there is little understanding about the impact of marketing strategies on foreign market entry-mode decisions or how mode structure influences a firm's ability to effectively implement their chosen strategy.

Top management team. As we know from decision-making research the characteristics of the decision maker or decision-making team play a critically important part in determining what choices are made (Wubben and Wangenheim, 2008). However, in the entry-mode literature managerial influences are seldom included. There are a few recent papers that have started to look at the decision-making process using real-options theory (Brouthers, Brouthers, and Werner, 2008b) and others that have examined manager characteristics (Brouthers and Hennart, 2007). But much more work needs to be done in this area.

For example, network theory has recently been used to help explain why managers make certain decisions (Rindfleisch and Moorman, 2001). Networks are the groups of individuals, both personal and professional, with whom the manager interacts, exchanges information, and shares resources. Researchers need to examine network connections, inside and outside the firm, to gain an understanding of how network relationships may influence the mode-choice decision. For example, do some networks discourage the use of specific entry-mode types or do some networks enhance the opportunity to use other mode types?

Organizational culture may also have a significant influence on the foreign market entry-mode choice decision. Firms may have cultural-based beliefs and values that eliminate certain mode choices from consideration or that influence the type of information that managers consider relevant to the mode-choice decision. Organizational cultural values can have an important impact on the decision-making process and outcome (Deshpande and Webster, 1989), yet entry-mode scholars have tended to ignore the impact of organizational culture in studies of entry-mode choice.

Furthermore, because foreign market expansion decisions are made at different levels of the firm, depending often on firm size, it is important that we obtain market entry decision data from the correct respondents. Studies looking at the impact of top managers in very large organizations, for example, may not be examining the right level of analysis. It is often the case in larger firms that SBU managers or product managers make these important expansion decisions. At present both marketing executives and middle managers have been conspicuously absent from entry-mode research. Obtaining data from these sources may be difficult but the value added could greatly increase our knowledge of how these decisions are really made.

As researchers in other areas have noted, managerial biases can have an important impact on the decisions they make. Those interested in foreign market expansion can help us gain

a better understanding of entry-mode choice by exploring the impact of factors that create biases in managerial decisions hence prohibiting managers from making the best entry structure decisions.

Conclusions

Although foreign market entry-mode scholarship has seen strong growth in recent years, there are still many unanswered questions. Even some of the most widely explored issues need further work as answers are still elusive as to the best way to make the entry-mode decision. Marketing practitioners and researchers can make an important contribution to our knowledge of market entry and expansion by focusing on how important marketing strategy issues and managerial biases influence these decisions.

In this article, we reviewed past research and highlighted several areas where new research is needed and where current research has failed us. By focusing on newer issues like market orientation, brand strategy, and product mix, researchers can help bring greater realism to our decision-making models. In addition, by gathering information from the appropriate level of manager in the firm and considering how issues like network relationships and organizational culture influence managerial decisions, future research can improve decision-making models that will help firms make better decisions in the future.

Finally, there are still enormous challenges ahead for researchers to gain a greater understanding of how market entry and expansion decisions are made. This expanded research can help facilitate better managerial practices and improve international performance for firms. Marketing scholars can have a substantial impact on this research by examining marketing issues and the influence of marketing managers in taking effective mode-choice decisions.

Bibliography

Anderson, E. and Coughlan, A.T. (1987) International market entry and expansion via independent or integrated channels of distribution. *Journal of Marketing*, **51**, 71–82.

Bello, D.C. and Lohtia, R. (1995) Export channel design: the use of foreign distributors and agents. *Journal of the Academy of Marketing Science*, **23** (2), 83–93.

Bradley, F. and Gannon, M. (2000) Does the firm's technology and marketing profile affect foreign market entry? *Journal of International Marketing*, **8** (4), 12–36.

Brouthers, K.D. (2002) Institutional, cultural and transaction cost influences on entry mode choice and performance. *Journal of International Business Studies*, **33** (2), 203–221.

Brouthers, K.D. and Brouthers, L.E. (2003) Why service and manufacturing entry mode choices differ: the influence of transaction cost factors, risk and trust. *Journal of Management Studies*, **40** (5), 1179–1204.

Brouthers, K.D., Brouthers, L.E., and Werner, S. (2008a) Resource-based advantages in an international context. *Journal of Management*, **34** (2), 189–217.

Brouthers, K.D., Brouthers, L.E., and Werner, S. (2008b) Real options, international entry mode choice and performance. *Journal of Management Studies*, **45** (5), 936–960.

Brouthers, K.D. and Hennart, J.F. (2007) Boundaries of the firm: insights from international entry mode research. *Journal of Management*, **33**, 395–425.

Canabal, A. and White, G.O. (2008) Entry mode research: past and future. *International Business Review*, **17**, 267–284.

Delios, A. and Henisz, W.J. (2000) Japanese firms' investment strategies in emerging economies. *Academy of Management Journal*, **43** (3), 305–323.

Deshpande, R. and Webster, F.E. (1989) Organizational culture and marketing: defining the research agenda. *Journal of Marketing*, **53**, 3–15.

Erramilli, M.K. and Rao, C.P. (1993) Service firms' international entry-mode choice: a modified transaction-cost analysis approach. *Journal of Marketing*, **57**, 19–38.

Geyskens, I., Steenkamp, J.B.E.M., and Kumar, N. (2006) Make, buy or ally: a transaction cost theory meta-analysis. *Academy of Management Journal*, **49** (3), 519–543.

Rindfleisch, A. and Moorman, C. (2001) The acquisition and utilization of information in new product alliances: a strength-of-ties perspective. *Journal of Marketing*, **65**, 1–18.

Sanchez-Peinado, E., Pla-Barber, J., and Hebert, L. (2007) Strategic variables that influence entry mode choice in service firms. *Journal of International Marketing*, **15** (1), 67–91.

Tihanyi, L., Griffith, D., and Russell, C.J. (2005) The effect of cultural distance on entry mode choice, international diversification and MNE performance: a meta-analysis. *Journal of International Business Studies*, **36** (3), 270–283.

Wubben, M. and Wangenheim, F.V. (2008) Instant customer base analysis: managerial heuristics often 'get it right'. *Journal of Marketing*, **72**, 82–93.

Zhao, H., Luo, Y., and Suh, T. (2004) Transaction cost determinants and ownership-based entry mode choice: a meta-analytical review. *Journal of International Business Studies*, **35** (6), 524–544.

standardization/adaptation of international marketing strategy

David A. Griffith

Introduction

The appropriateness of standardization/adaptation of international marketing strategy has been a topic of importance to firms since the earliest times of international expansion. Most credit the stimulation of this debate to the practitioner realm in 1923 where a difference appeared between the approach to advertising by Goodyear's David Brown and Bausch & Lomb's Carl Propson. Brown (1923) viewed humanity as possessing common attributes, thus allowing for standardization. On the other hand, Propson (1923) argued that adaptation was often necessary to appeal to divergent local markets. Both approaches had their proponents and detractors with anecdotal evidence supportive of each.

The initiation of academic study on this important issue began in the late 1950s (Pratt, 1956), finding strong academic inquiry in the 1960s (Donnelly and Ryans, 1969), as the academic marketing discipline moved toward a more theoretical mode of inquiry. Since the initiation of academic inquiry on this topic, research on the topic has continued to proliferate and remains unabated (Donnelly and Ryans, 1969; Harvey, 1993; Szymanski, Sundar, and Varadarajan 1993; Onkvisit and Shaw, 1999; Griffith, Hu, and Ryans, 2000; Zou and Cavusgil, 2002; Ryans, Griffith, and White, 2003; Katsikeas, Samiee, and Theodosiou, 2006; Shoham *et al.*, 2008). Of the research that has been advanced, the vast majority of research presented prior to 1990 was conceptual/theoretical/anecdotal in nature. However, a significant number of empirical investigations have been put forth over the last 20 years (Cavusgil and Zou, 1994; Katsikeas, Samiee, and Theodosiou, 2006) that have provided a foundation for further clarification of this topic via the creation of contingency models.

Defining Standardization/Adaptation

While a significant number of studies have been conducted on this important topic, an examination of a selection of seminal articles in the field demonstrates considerable divergence in foundational definitions. One such definitional concern relates to standardization/adaptation. For example, Jain (1989) defined *standardization* as a common marketing program on a worldwide basis. Szymanski, Sundar, and Varadarajan, (1993, p. 1) view *marketing-strategy standardization* as the "standardization of the pattern of resource allocation among marketing-mix variables across national markets." Alternatively, Cavusgil and Zou (1994) viewed the issue under the general "degree" of adaptation. Further, in integrating the standardization, configuration-coordination, and integration views, Zou and Cavusgil (2002, p. 42–43) conceptualize *global marketing strategy* (GMS) (*see* GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES; GLOBAL MARKETING STRATEGY) as the degree to which a firm globalizes its marketing behaviors in various countries through standardization of the marketing-mix variables, concentration and coordination of marketing activities, and integration of competitive moves across markets (*see* MARKETING STRATEGY IMPLEMENTATION; MARKETING MIX). Clearly, a consensus eludes researchers in the definition of this important topic. Lack of a consistent perspective on what constitutes marketing-strategy standardization/adaptation has created a fundamental limitation for advancement of research in this area.

Further inconsistencies can be viewed in relation to the definitional aspects of the outcomes of marketing-strategy standardization/adaptation. Specifically, the effectiveness of a firm's international marketing-strategy standardization/adaptation efforts is measured in terms of one performance aspect or another. For instance, Jain (1989, p. 76) argued that "standardization should be based on economic payoff, which includes financial performance, competitive advantage, and other aspects." Others, such as Cavusgil and Zou (1994), conceptualize performance as the extent to which a firm's objectives, both economic and

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

strategic, are achieved through planning and execution. Others, taking a more consumer orientation (*see* CUSTOMER RELATIONSHIP MANAGEMENT), explored the acceptability to consumers of product attributes, messages (*see* INTERNATIONAL ADVERTISING – IS THERE STILL A STANDARDIZATION VERSUS LOCAL ADAPTATION DEBATE?), brands (*see* BRAND STRATEGY; GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES), and so on. Still others, sidestep the issue of effectiveness by not conceptually defining it, but rather assuming it to be generally understood.

Given the diversity of conceptual definitions of both international marketing-strategy standardization/adaptation and the lack of consistency related to the conceptualization of the outcome effectiveness of marketing-strategy standardization/adaptation, it is not surprising that inconsistencies have been found within the literature relating to this issue (Onkvisit and Shaw, 1999; Zou and Cavusgil, 2002). This article does not work to resolve these conceptual distinctions but rather to address how differences in approaches to these distinctions have stimulated work in four key areas that are central to the understanding of the standardization/adaptation of marketing strategy: (i) type of standardization/adaptation, (ii) standardization/adaptation unit of analysis, (iii) market homogeneity and the effectiveness/efficiency argument, and (iv) models of standardization/adaptation.

Types of Standardization/Adaptation

Within both practice and research, standardization/adaptation of marketing strategy centers on two fundamental types: program and process (Sorenson and Wiechmann, 1975; Jain, 1989). Program standardization involves the tactical elements of marketing-strategy employment and is principally founded in the marketing mix, for example, product, place, promotion, and price (cf., Cavusgil and Zou, 1994). As such, program standardization/adaptation examines whether specific elements of the marketing mix are consistent across markets. For example, differences in electric voltage for a product or the color used on promotional elements across the United States and Argentina would be characteristic of program adaptations. Thus, a program approach to marketing-strategy standardization/adaptation looks to the inherent tactical execution elements of the firm's marketing efforts (such as pricing) across markets. Alternatively, process standardization involves the development of a common approach to the procedures underlying marketing-strategy activities, such as establishing (and evaluating) a distributor network (*see* SUPPLY CHAIN MANAGEMENT STRATEGY) and translating currency equivalents necessary for pricing (*see* INTERNATIONAL PRICING OBJECTIVES AND STRATEGIES) (Jain, 1989). A standardized process for, say, developing relationships is quite different from a standardized program (i.e., specific management tactics to be used to administer the relationship) providing insights into the manner in which marketing strategy is developed and managed (cf., Griffith, Hu, and Ryans, 2000). As such, process standardization/adaptation incorporates the procedural elements underlying the aspects of marketing strategy and tactics (*see* MARKETING STRATEGY; MARKETING STRATEGY IMPLEMENTATION).

While both program and process elements of marketing strategy are important to understanding a firm's success, the fact that program elements of a firm's marketing mix are more clearly observable and measurable has lead to a research dominance on this aspect of marketing strategy when compared to process elements (which are typically not easily observable and highly variant across firms). Future research in this area will continue to address program and process issues (and hopefully the related nature of process and program standardization/adaptation) to gain a deeper understanding of this area.

Standardization/Adaptation Unit of Analysis

The unit of analysis of standardization/adaptation research has also been divergent. The unit of analysis specifies whether marketing-strategy standardization/adaptation is examined at a holistic marketing-strategy level, the individual marketing-mix element or

process level, or whether specific subelements within one of the marketing-mix elements or specific subprocesses of a marketing-strategy procedure ara examined.

At the most aggregate level, marketing strategy is viewed holistically with all elements of the marketing strategy (*see* MARKETING STRATEGY MODELS; MARKETING STRATEGY) aggregated at the firm level (whether program or process) for the assessment of the firm's approach to marketing-strategy standardization/adaptation. For example, Zou and Cavusgil (2002) conceptualize the construct of GMS (*see* GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES) by aggregating across the firm's marketing activities working to understand, at the highest level, the influence of marketing strategy on firm performance issues. Current state-of-the-art findings related to this unit of aggregation are presented in a meta-analysis conducted by Katsikeas, Samiee, and Theodosiou (2006). Katsikeas, Samiee, and Theodosiou (2006) note that when observed at this level, stable performance implications are only obtainable when contingent factors are accounted for in the model.

The second unit of analysis investigated is the individual program or process related to a specific element of the marketing mix, for example, advertising strategy (*see* INTERNATIONAL ADVERTISING – IS THERE STILL A STANDARDIZATION VERSUS LOCAL ADAPTATION DEBATE?), standardization/adaptation as a program element, or the procedures for developing advertising strategy as a process element. Researchers have often examined the individual marketing-mix elements and the degree of standardization/adaptation engaged in by the firm in respect of each marketing-mix element, such as promotion standardization/adaptation or channel standardization/adaptation. For example, Shoham *et al.* (2008) studied the behavioral and performance influences of international standardization of channel management (*see* INTERNATIONAL MARKETING CHANNELS), thus providing insights to the process of distribution management. Research at this unit of analysis, when investigating program elements, often aggregates aspects of the firm's marketing mix (e.g., product focus research investigating labeling, packaging, ingredients, etc.).

Lastly, researchers have argued for the need for fine-grained insights into how individual elements (either program or process) are employed. These scholars argue that practicing marketing managers employ the lowest level of analysis (i.e., decisions are made over every element of the firm product, inclusive of brand name, product specifications, packaging color, and size, etc.) and, as such, only through a detailed examination at the lowest level of analysis can a strong understanding of the implications and appropriateness of standardization/adaptation of marketing strategy be gained. For example, Griffith, Chandra, and Ryans (2003) explore individual elements of the promotional mix demonstrating unique performance implications in relation to each.

The extant research in this area demonstrates that unique insights can be gained from exploring each level. The future of standardization/adaptation research will continue to focus on all three units of analysis bringing forth new insights for academic inquiry and practical application.

Market Homogeneity and the Effectiveness/Efficiency Argument

The issue of the effectiveness versus the efficiency of the standardization/adaptation of marketing strategy resides in whether there is consumer homogeneity and/or the movement toward homogeneity (Donnelly and Ryans, 1969; Levitt, 1983; Walters, (1986); Samiee and Roth, 1992; Szymanski, Sundar, and Varadarajan, 1993; Zou and Cavusgil, 2002) (*see* CROSS-CULTURAL PSYCHOLOGY OF CONSUMER BEHAVIOR; SOCIETY, CULTURE, AND GLOBAL CONSUMER CULTURE). Those researchers who view markets, or consumer wants and needs, as being homogeneous argue that the standardization of marketing strategy is more effective and efficient as it allows for the lowering of costs, via economies of scale, and thus increasing margins for a firm (Levitt, 1983; Jain, 1989) while simultaneously meeting the needs of the market. Alternatively, those who view markets as being heterogeneous, and therefore containing consumers with differing consumer wants and needs, perceive greater value delivery via adaptation of marketing-strategy elements

(Donnelly and Ryans, 1969; Harvey, 1993), (the perspective here is that since consumer wants and needs are not homogeneous, efficiency under the standardization approach is an illusion). Despite its centrality to the focal research issue, the extent of homogeneity of markets for marketing-strategy standardization has been given little attention within the literature (this issue has primarily been the research domain of market segmentation without strong relation to the standardization/adaptation implications). Rather than focusing attention on market homogeneity directly, researchers have instead explored a number of consumer response elements believed to be important to GMS cross-culturally, such as brand loyalty (*see* CONSUMER BRAND LOYALTY), risk perception, and, most recently, brand personality. Given the disparity of consumer response in these studies, it appears as though the empirical evidence would contradict the argument for homogeneity of markets. However, it is important to note that firms often target very select markets (*see* MARKET SEGMENTATION AND TARGETING) and, therefore, consumer homogeneity often does (or can) exist at some level for the firms when a product is offered globally.

With underlying heterogeneity and homogeneity across markets, at least in some respect, the issue of efficiency and effectiveness of the standardization/adaptation of marketing strategy is evident. When homogeneity of markets exists, the standardization of international marketing strategy reduces costs by spreading costs over a greater number of markets, thus reducing average costs; the theoretical linkage is that lower costs achieved through standardization increases efficiencies while retaining the same level of "effectiveness" (where effectiveness relates to consumer response). Alternatively, where consumer heterogeneity exists, proponents of adaptation argue that international marketing-strategy adaptation that meets local consumer needs and wants is more effective (Harvey, 1993). While empirical support exists for the relationship between adaptation and performance, the issue of effectiveness of the adaptation of marketing strategy remains. For instance, could an adapted marketing strategy be effective, yet not increase performance? Of course, it could depend upon the measures of effectiveness and performance one selects. A marketing strategy can have many differing objectives, for instance, increase brand awareness. The direct linkage between adaptation and performance may not be comprehensive – or at the least, fully explanatory (a recent meta-analysis by Katsikeas, Samiee, and Theodosiou ((2006)) provides substantive evidence for the need for a contingent approach to understanding the standardization/adaptation performance relationship).

Thus, when taken together, we can see that the effectiveness or efficiency of the standardization/adaptation of marketing-strategy decision is founded on consumer homogeneity/heterogeneity. With increased globalization (*see* SOCIETY, CULTURE, AND GLOBAL CONSUMER CULTURE; GLOBAL CONSUMERISM AND CONSUMPTION), it is argued that some customer segments will become more homogeneous allowing for greater effectiveness and efficiencies offered by standardization. However, until that time, adaptation to local market needs fulfills the central position of marketing, which is value delivery to local needs.

Models of Standardization/Adaptation

Current models of standardization/adaptation recognize the importance of the firm's operating environment through the employment of contingency models, that is, marketing strategy is a strategic response to competitive conditions founded upon the firm's resources and its environment. Building on this conceptual foundation, researchers proffer the appropriateness of employing a strategy–environmental coalignment theoretical perspective. Under this theoretical perspective, marketing strategy is viewed as a strategy–environment fit maximization exercise, where when a firm's strategy is aligned with the firm's internal and external factors, a firm performs at an optimal level, thus achieving above normal returns (*see* FORCES AFFECTING GLOBAL INTEGRATION AND GLOBAL MARKETING). This conceptual framework incorporates, at

a minimum, three key features critical to understanding the inherent marketing strategy–performance relationship, that is, (i) the unit of analysis underlying the framework is the individual venture, (ii) the framework posits that firm/brand performance involves multiple outcomes, and (iii) the framework is presented in general terms, thus allowing for flexibility in the internal and external variables included, thus allowing for industry-specific deviations in salient criteria.

First, in terms of the unit of analysis of the venture, researchers contend that it is centrally important to the scope of responsibility of marketing strategy. If the scope of responsibility is at the brand level, then brand-level criteria should be employed; if at the strategic business unit level then analysis should occur at that level. Only when the specified unit of the venture can be identified will effective understanding of marketing-strategy standardization/adaptation be observable. This issue is becoming increasingly problematic where primary data from a specific venture of the firm (e.g., brand) is captured and matched with firm-level performance data (when the firm has multiple brands). Although this approach works to overcome issues of common method bias, the violation of the scope of marketing-activity responsibility coupled with differences in units of measurement (a single brand's marketing-strategy decisions paired with the firm's performance – which included multiple brands) results in bias.

Second, the framework posits that multiple performance outcomes be considered. In the marketing-strategy literature, research continues to demonstrate differential performance results from a single antecedent element when performance metrics are varied. As such, decomposing performance metrics into aspects, such as financial, strategic, or behavioral, becomes important for clarifying the effects of marketing-strategy standardization/adaptation on firm performance (*see* MARKETING METRICS). For example, a firm may adapt its marketing strategy within a new market to serve local needs. By doing so, the long-term strategic goal of gaining market share could be achieved, while financial performance may denote lower margins (due to increased marketing costs associated with the adaptation).

Third, both internal and external factors are conceptualized to influence marketing-strategy standardization/adaptation (Cavusgil and Zou, 1994). Internal factors relate to the firm and encompass a wide variety of issues, inclusive of, but not limited to, management forces (such as international experience of management, number of languages spoken by key managers, specific country/regional interests of key managers) as well as firm-related factors (such as the firm's involvement in international business activities, the transferability of the firm's current offering, financial position of the firm, the firm's access to capital in new markets). External factors relate to the operating environment of the firm encompassing both competitive and market characteristics. For example, issues such as the competitive intensity in the market, legal requirements for market entry, level of consumer demand, level of consumer sophistication, and so on, are all considerations determining marketing-strategy standardization/adaptation.

Conclusion

The issue of the standardization/adaptation of marketing strategy has made great progress; however, much is yet to be done. While the field has identified the research question, it is yet to come to a consensus on the conceptual domains of the central elements. The field will struggle with this issue as it moves forward. The four issues discussed here are less points of argument than points of clarification, wherein, when addressed, clearly identifiable solutions are evident. The four issues, when taken into consideration, provide a clearly focused research avenue under a contingent-model approach for the investigation of the standardization/adaptation of marketing strategy.

Bibliography

Brown, D.L. (1923) *Export Advertising*, Ronald Press, New York.

Cavusgil, S.T. and Zou, S. (1994) Marketing strategy-performance relationship: an investigation of the empirical link in export market venture. *Journal of Marketing*, **58** (1), 1–21.

Donnelly, J.H. Jr. and Ryans, J.K. Jr. (1969) Standardized global advertising: a call yet unanswered. *Journal of Marketing*, **33** (2), 57–64.

Griffith, D.A., Chandra, A., and Ryans, J.K. Jr. (2003) Examining the intricacies of promotion standardization: factors influencing advertising message and packaging. *Journal of International Marketing*, **11** (3), 30–47.

Griffith, D.A., Hu, M.Y., and Ryans, J.K. Jr. (2000) Process standardization across intra- and inter-cultural relationships. *Journal of International Business Studies*, **31** (2), 303–324.

Harvey, M.G. (1993) Point of view: a model to determine standardization of the advertising process in international markets. *Journal of Advertising Research*, **33** (4), 57–64.

Jain, S.C. (1989) Standardization of international marketing strategy: some hypotheses. *Journal of Marketing*, **53** (1), 70–79.

Katsikeas, C.S., Samiee, S., and Theodosiou, M. (2006) Strategy fit and performance consequences of international marketing standardization. *Strategic Management Journal*, **27** (9), 867–890.

Levitt, T. (1983) The globalization of markets. *Harvard Business Review*, **61** (3), 92–102.

Onkvisit, S. and Shaw, J.J. (1999) Standardized international advertising: some research issues and implications. *Journal of Advertising Research*, **39** (6), 19–24.

Pratt, E.E. (1956) Building export sales-advertising. *The International Advertiser*, **1** (8), 19–22.

Propson, C.F. (1923) Illustrating the foreign campaign, *Export Advertising Practice*, Prentice-Hall, New York.

Ryans, J.K. Jr., Griffith, D.A., and White, D.S. (2003) Viewpoint: standardization/adaptation of international marketing strategy: necessary conditions for the advancement of knowledge. *International Marketing Review*, **20** (6), 588–603.

Samiee, S. and Roth, K. (1992) The influence of global marketing standardization on performance. *Journal of Marketing*, **56** (2), 1–17.

Shoham, A., Brencic, M.M., Virant, V., and Ruvio, A. (2008) International standardization of channel management and its behavioral and performance outcomes. *Journal of International Marketing*, **16** (2), 120–151.

Sorenson, R.Z. and Wiechmann, U.E. (1975) To what extent should a consumer goods multinational vary its marketing from country to country? *Harvard Business Review*, **53** (3), 38–54.

Szymanski, D.M., Sundar, G.B., and Varadarajan, P.R. (1993) Standardization versus adaptation of international marketing strategy: an empirical investigation. *Journal of Marketing*, **57** (4), 1–14.

Walters, P.G. (1986) International marketing policy: a discussion of the standardization constructs and its relevance for corporate policy. *Journal of International Business Studies*, **17** (2), 55–69.

Zou, S. and Cavusgil, S.T. (2002) The GMS: a broad conceptualization of global marketing strategy and its effect on firm performance. *Journal of Marketing*, **66** (4), 40–56.

international product innovation and development

Roger J. Calantone and Janell D. Townsend

For millennia, individuals and firms have tried to create local or regional monopolies through differentiation. By offering a good, product, or service that was better in some (generally tangible) manner the aspiring monopolist could command a price differential from the price paid for the ordinary good purchased by the "mass" market. When unobservable quality or exclusivity was the differentiator, special "trademarks" denoted the distinctive outputs of the maker. As countries industrialized various economic sectors in the 1800's efficiency allowed prices to fall dramatically and mass production with little differentiation began to dominate the scene. In the 1830's clockmakers in Connecticut were able to produce clocks so cheaply that almost every working household had the means to purchase a mantle place clock, whereas in Europe at that time, clocks were still made individually – only the wealthy had clocks. Skilled workers in industries such as this, possessing both the explicit and tacit learning of their employer's processes, immigrated to new countries where they became agents of innovation in processes that increased productivity. The most common base differentiator was price accomplished through cost economies of process reengineering allowing for the mass diffusion of innovation across markets. Such engineering of efficiency reduced waste, which coincidentally has beneficial effects on quality delivered. No one country was a single source of this systematic industrialization, although resource endowments allowed some to move process innovation forward more rapidly. Thus, process innovation delivered production economies to mass markets, and in today's global marketplace, differentiation is derived from product innovation itself.

This article reviews international product innovation taking an activist firm perspective with respect to global product development. First, an overview of the nature of strategic innovation management in global markets is presented. The strategic intent and role of standardization versus adaptation in global products is then discussed. Next follows a delineation of significant factors associated with organizing and managing global product innovation, and a conclusion summarizes.

Strategic Innovation for Global Markets

In today's global marketplace, the process of innovation relies heavily on flexibility, speed, and efficiency as the rates of technological innovation have increasingly shortened product life cycles, and enabled a broader and more diverse set of competitors. During this period of increasing resource constraints as well as greater competitive threats, companies are faced with the need to accelerate product development (Rothwell, 1994). The growing complexity and pace of industrial technological change are forcing firms to first understand the role and importance of global product innovation, how this fits with the firm's level and strategic orientation toward globalization, and the interaction of these forces with the international marketing concept of the company.

New product development (NPD) involves the necessary but competing goals of minimizing risk by acquiring sufficient market information while reducing costs and time to market, thus escalating the importance of NPD process design and implementation (Harmancioglu *et al.*, 2007). In other words, a firm's NPD processes and how they are implemented are vital for decreasing lead time and increasing innovation productivity. NPD processes involve a series of stages aimed at delivering a functional commercial benefit to customers (Calantone and DiBenedetto, 1995). Proficiency in executing NPD processes is important because it determines the degree to which businesses can meet and/or exceed customer demands, and thus succeed in a global marketplace (Cooper, 1991). The stages of NPD are rather universal, although various authors break it into as few as three steps and as many as forty, according to a particular application. There is no "one size fits all" approach to NPD, so the organizational design elements of each firm are different and come into play at different stages of the NPD process. However, each element plays a

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

vital role, requiring careful consideration and planning.

Global Orientation of Innovation

Globalization of the firm has been conceptualized as the transformation of businesses from domestic, to multinational, to those with global scale and scope (Ghoshal, 1987; Perlmutter, 1969). Following the basic tenets of incremental internationalization applied to the globalization process (Johanson and Vahlne, 1977), increasing knowledge of global markets yields a hierarchical structure of potential structural orientations. This is based on complex interactions with respect to the role of market attractiveness, experiential learning, and mimetic behavior in globalization patterns (Townsend, Yeniyurt, and Talay, 2009). The level of commitment is not static over time, but dynamically ranges from initial market entry approaches such as casual exporting to very high levels of integration for the more globally oriented firms. Identifying the current firm orientation toward internationalization provides a basis for understanding the underlying philosophy that guides the organization's approach toward international strategy and decision-making (Cateora and Graham, 2009). Global orientation is conceptualized as the organization's ability to view the entire world as its marketplace, not relying on individual markets or regions exclusively, or independently. While a multidomestic company treats customers and competitors in each country or region on a stand-alone basis, a global company takes an integrated approach across countries and regions (Birkinshaw, Morrison, and Hulland, 1995; Zou and Cavusgil, 2002). This means emphasizing the global success of the firm, as opposed to accentuating nation- or market-based measures (Ohmae, 1989), and is consistent with Perlmutter's (1969) original conceptualization of the geocentric firm. Yet, the implementation of a global orientation remains a major challenge for leadership, when trying to integrate a global strategy and a global structure (Roth, Schweiger, and Morrison, 1991; Yip, 1992; Zou and Cavusgil, 2002). High-level strategic orientations such as these impact all elements of the marketing mix, with innovation and product development being among the core processes undertaken by the firm (Townsend *et al.*, 2004).

Three general international marketing concepts are commonly used to define level of commitment: domestic market extension concept, multidomestic market concept, and global marketing concept (Cateora and Graham, 2009). The domestic market extension concept is characterized by the selling of domestic product in foreign markets. Minimal to no adaptation is done to the marketing mix. The multidomestic market concept is characterized by a firm's recognition of the importance of the differences in foreign markets along with the recognition of the importance of international business to its operations. Companies that operate under this approach tend to look at foreign markets as being vastly different, and accordingly, requiring a unique strategy for each country. At the core of the global marketing concept are the efficiencies that can be obtained through standardization. In general, strategy is set at a global level with the understanding that some decisions are affected by local influences and will need to be looked at on a country-by-country, or region-by-region basis. From this conceptual perspective, the entire world is seen as a market, with segments that span multiple countries.

Product Policy: Standardization Versus Adaptation

The globalization of product marketing originates from the debate about the relative level of marketing mix standardization (Buzzell, 1968). It appears that consumer expectations around the world are beginning to converge in terms of needs and expectations as products that deliver a consistent identity have become more viable. A study identifying "marketing universals" found that there are few differences in consumers' use of quality signals across cultures – yet, only for selected levels of segmentation (Dawar and Parker, 1994). Research does suggest a general degree of homogeneity across market segments which transcend national boundaries (Yavas, Verhage, and Green, 1992). This is supported by recent findings that socioeconomic variables moderate the effects of cultural dimensions on the acceptance of new products (Yeniyurt and

Townsend, 2003), and the degree of foreignness of new products is having less of an effect on performance over time (Townsend, Calantone, and Schmidt, 2003). Thus, product policy related to standardization versus adaptation is the function of the firm's strategic orientation coupled with the degree of homogeneity across geographic and cultural markets.

A firm's international product policy is critical due to cost ramifications, and the inimical prospect that value creation and transmittal manifests in the product. Packaging style, quality, labeling, and brand name may seem trivial to some. Yet, these characteristics play a major part in international marketing due to the degree of calibration with cultural norms and preferences. Some products may need to be slightly altered and others not at all. Observations from the marketplace seem to support the idea of finding an appropriate balance between standardization and adaptation (Cavusgil, Zou, and Naidu, 1993; Jain, 1989b), with the premise being to embrace the concept of being global, but acting locally as necessary.

Standardization. A standardized product policy generally means that the firm will create a standard product to be sold in all markets served. However, companies will sometimes market their current domestic product internationally, as is, under the same brand name, in the same packaging, and with the same level of quality. The product policy does not change irrespective of the target market. While this approach preserves the low cost producer idea, long production runs, undifferentiated marketing, and economies of scale and experience, driving per unit variable and fixed costs downward, it ignores an inherent need for variety within any culture, and the differences of tastes between cultures. This is hubris in the face of the diversity of other cultures, and usually market punishment is quick and sure.

The primary benefits of a standardization approach to product development are the production economies and other cost savings that can be obtained. Supporters of standardization believe that price, quality, and reliability will offset any differential advantage that having a culturally adapted product would provide in the eyes of the customer (Jain, 1989b). A standardized product policy can be useful because economies of scale are created in activities, especially in research, development, manufacture, and marketing (Kuvykaite, 2008). Some market segments are the same no matter where they exist geographically (Katz, 1987). Proponents of standardization argue that with the increased levels of global communication and other worldwide socializing, the tastes, needs, and values in a significant sector of the population across all cultures has become more homogeneous. The argument is that market segmentation is based on the lifestyle of the consumer, and standardized products can be marketed globally when the segmentation scheme is done using criteria other than geography alone. Product standardization is a forerunner of overall marketing mix standardization, reducing the complexity of operations. Standardization allows for less complex organizations that are easier to manage and control (Majaro, 1982).

No policy is without disadvantages though. Marketing flexibility is lost because of the inability to match the product to local requirements. Standardization suppresses entrepreneurship because a standard global product is accepted in all markets, complacency sets in and fresh new ideas are few and far between (Wind, 1986) – some personnel may be lost to organizations that provide more opportunities for creativity in marketing and product design (Majaro, 1982). Also, standardized products can be too complicated for some markets and too simple for other markets; some markets may need extensive training before accepting a product, while others may find the product too simple and will thus reject it (Wind, 1986).

Industrial customers around the world are generally more similar than their consumer goods counterparts because their purchasing decisions are driven less by attitudes and feelings, and more by economic considerations. Because of this, standardization is typically seen as the strategy of choice for manufacturers of industrial products. The main concerns of industrial customers are service, dependability, quality, performance, and cost (Cateora and Graham, 2009). Also, in recent years, there has been a trend toward more international standards (e.g., ISO standards) (Usunier, 2003), providing

impetus for using a standardization strategy for industrial products.

Adaptation. Adapting products for international markets simply means expanding the organization's product line (Calantone, Cavusgil, and Schmidt, 2004). Supporters of adaptation say it is inevitable. The most important objective of a firm is not minimization of costs through standardization, but long-term profitability, achieved by satisfying various consumer needs in different countries, thus ensuring greater sales (Pimblett, 1997). Many of the benefits of an adaptive product policy are obvious. For instance, the more a product is tailor-made for a specific market, the better it will fit the needs of the customers (Calantone, Cavusgil, and Schmidt, 2004). This, in turn, should lead to higher sales and sustained growth. A product adapted to a target market based on market research is more likely to succeed, and therefore carries less inherent risk than a standardized product.

Drawbacks to the adaptation approach can include increased costs related to research and development and the loss of scale economies. There may also be an increase in the complexity of the organization in response to the addition of foreign market operations to the preexisting domestic market operations. This will add a level of complexity to the management and an overall control of the organization. When defining the level of international commitment, management should ensure that they have the appropriate level of resources committed to the foreign endeavors (Cateora and Graham, 2009). From a consumer's standpoint, multiple products with different packaging and different brand images can cause identity or credibility problems.

Adaptations can be grouped into two categories: obligatory adaptation and discretionary adaptation. Obligatory adaptations are defined as those that an exporter is forced to undertake because of regulations that must be met in order to enter a foreign market or because of external environmental factors (e.g., climate considerations). Discretionary adaptations are voluntary adaptations that a firm undertakes in order to better align its product with market needs or other cultural factors (Jain, 1989a).

Several considerations come into play when determining the level of adaptation necessary for a product in a foreign market. In order to understand all the possible ways a product can be adapted it should be deconstructed into components based on benefits delivered. Major adjustments to the core component can be costly if changes to the production processes must be made to accommodate the specialized products. This may require a large capital investment. Auxiliary components include things such as packaging which protect the product's integrity, but also serve as a communications platform, sometimes tightly regulated by governmental edits.

The importance of the features contained within the packaging component depends on the need that the product is designed to serve. For example, in countries where literacy levels are relatively low, packaging must include symbols or pictures to aid the consumer in identifying the contents of the package and the appropriate usage. In other instances there may be legal requirements for labeling (e.g., information printed in multiple languages). It could also be the case that package sizes are regulated by law. External environment factors, such as humidity, could also bring about the need for adjustments to packaging. In some countries, such as Japan, the quality of the packaging has a direct impact on the consumer's perception of the quality of the product within.

In addition to the physical and service aspects of product adaptation, the impact of the symbolic attributes related to a product must also be examined (Usunier, 2003). In order to determine the symbolic attributes a product may have, a firm needs to first understand the culture of the country in which the product will be sold, including elements such as materialism, social institutions, belief systems, and language. There are two types of cultural knowledge that are necessary: factual and interpretive. Factual knowledge can be easily learned; interpretive knowledge, conversely, requires cultural insight usually acquired through personal experience.

Organizational Design

Organizational design elements are critical to success if product innovation and management

is to be successful across global markets. Global organizations need to determine and achieve a balance between central authority and responsiveness to local preferences that optimizes their business position (Johansson and Yip, 1994; Roth, Schweiger, and Morrison, 1991). Influential organizational design elements include formally planned stages, senior level involvement, business case preparation, customer input, and cross-functional integration (Barczak and McDonough, 2003), while a business case delineates project goals, market projections, and possible product specifications (Harmancioglu *et al.*, 2007). Coordination mechanisms in NDP include linking electronically geographically dispersed parts of the organization via intranets, extranets, and so on (Boudreau *et al.*, 1998), best practice repositories, and lead centers of excellence (Frost, Birkinshaw, and Ensign, 2002).

Subsidiary integration and global product mandates. A general trend has been observed such that multinational corporations have begun initiatives focused on integrating value-added activities which were once globally dispersed. This global dispersion occurred as a response to host government import/export regulations and tariffs, but with the globalization of business in recent years, these types of dispersed organizational structures are no longer necessary. Utilizing formal and informal interfunctional coordination mechanisms allows organizations to achieve global responsiveness while balancing flexibility and efficiency (e.g., Bartlett and Ghoshal, 1987; Martinez and Jarillo, 1991). With increased globalization foreign subsidiaries are now being used in more specialized roles with greater market scope (e.g., exporting) but narrow functional and/or product responsibility (Birkinshaw, 2002). World product mandate gives global responsibility to a subsidiary for development, manufacturing, and marketing of a single product line. Although full-scope mandates of this nature are relatively rare, regardless of scope, the primary outcome of the mandate process is greater specialization in terms of focused product responsibility (Birkinshaw, 2002).

In terms of specialization there are two theoretical approaches: rationalization-integration and world product mandate. Rationalization-integration occurs when a subsidiary produces a component under assignment from the parent organization for the firm as a whole. Exporting is controlled by the subsidiary but upstream responsibilities such as development and design are controlled by the parent organization. Full-scope world product mandate, as mentioned previously, gives full control of development, manufacturing, and export marketing to the subsidiary. In this type of relationship, the subsidiary acts more as a partner than a subordinate to the parent and has a higher level of autonomy than in the rationalization-integration approach (Birkinshaw, 1996). In practice, a hybrid approach is more commonly observed; for example, a subsidiary may have global production and marketing responsibilities but utilizes central R&D resources for new product development.

There are four motives that are generally accepted classifications of subsidiary mandates: market-seeking, resource-seeking, efficiency-seeking, strategic asset-seeking (Birkinshaw, 1996). Each has a set of characteristics related to the business benefit the parent organization is attempting to achieve via the mandate. There are several challenges related to the establishment and management of subsidiary mandates. One such challenge is the restructuring of the organization to accommodate a new decentralized decision-making and reporting structure. The estimated value addition from the subsidiary should be able to cover the costs associated with this restructuring. Also, typically, the products assigned as a part of the subsidiary mandate approach are usually products at the end of the product life cycle. Care must be taken to ensure that the subsidiaries remain relevant to the current strategic vision of the parent organization even if the primary focus is on a product that is not at the forefront for management. Lastly, because of the specialized nature of these mandates, foreign subsidiaries are vulnerable to changes in the marketplace. If subsidiaries are unable to adapt to the market changes, or if organizations are unprepared to shift mandates to different subsidiaries to meet market needs, the mandate approach will be unsuccessful (Birkinshaw, 1996).

Open innovation. Traditionally, the ideas and concepts that feed innovation have been generated via experts and/or scientists within internal research and development departments. Recently, more organizations have adopted an approach which includes "open innovation" in the new product development cycle. Open innovation utilizes ideas and inspiration from "creative consumers." These creative consumers differ from mainstream consumers in that they are excited by new ideas whereas mainstream consumers tend to like what they already know. The open innovation theory proposes that tapping into these creative consumers will help to overcome the thinking that most market research is backward looking as opposed to the forward-looking approach that is needed for product innovation (Clegg, 2008).

The engaged consumer has always existed, but now they are easier to identify and access via social networking and user-generated communities. The openness emerging from user-generated forums is where the true gain accrues relative to the classic opinion/idea collection methods of surveys and focus groups. The use of the web as a means of communication gives companies access to consumers on a global scale which is more difficult and expensive using the classic methods of data collection. This broad and global perspective on consumer ideas gives companies an advantage when attempting to generate breakthrough innovations.

Although disruptive product innovation appears to be key to the long-term health of an organization, there is no assurance that the ideas generated from open innovation mechanisms ultimately lead to these highly coveted product outcomes. The volume of information acquired can itself hinder the creative process. Automated tools provide an information capture mechanism, but the screening and sifting task to discover something commercially successful can be frustrated in many ways. This uncertainty drives many firms to opt for incremental product line extensions that utilize their current business capabilities as opposed to the more risky breakthrough innovation route.

Cooperation in the new product development process. The complexities of the global marketplace have required companies to forge new vertical and horizontal alliances and to seek greater flexibility and efficiency in responding to market changes. These multifaceted and complex organizational relationships seek to establish or extend a firm's differentiation by way of an alliance, either vertically in its value chain or horizontally through either competitors or complementary companies. Since alliances allow for the pooling of resources, it stands to reason that they would create a broader range of resource opportunities in the product innovation process. Through alignment and extension, collaboration with a partner provides an opportunity to fulfill the requirements of a sustained competitive advantage, which cannot be achieved independently; through the efficient use of a partner's existing resources, the boundaries of the firm can be effectively extended. This includes knowledge sets that are both externally facing like culture and markets, and those that are internally oriented like product-specific processes.

In recent years, the trend has been for organizations to cooperate with different external partners as a way to enhance the efficiency and effectiveness of the new product development process, cut costs, and to reduce risk. These partners can include distributors, consumers, universities and research centers, and even competitors. Studies suggest that there is a positive relationship between cooperation and the achievement of success in the process of innovation. Cooperative alliances can be divided into two categories: (1) those based on synergies and complementary assets; (2) those based on growth opportunities and market power (Arranze and Arroyabe, 2008). Cooperation can be further identified as "vertical" or "horizontal" cooperation, respectively. Vertical cooperation (also known as supply chain cooperation) plays an important role in the collection of information on technologies, user needs, and markets. Partnerships with suppliers are seen as a complement to internal R&D activities as opposed to a substitute for them, and partnering with customers reduces the risks associated with market introduction. With horizontal cooperation, competitors may have complementary resources which will allow both parties to reduce costs and risks in large projects.

These types of partnerships are best suited for scenarios where either a strong common interest has been identified, for example, cooperating on the development of a new range of product or services, or scenarios where the resulting research leads to generic results (Arranze and Arroyabe, 2008).

There are a number of benefits derived from using a cooperative approach to new product development (Vilaseca-Requena, Torrent- Sellens, and Jimenez-Zarco, 2007) – for example, the establishment of work teams made up of experts in different functional fields who adopt flat structures (e.g., minimal layers between employees and management) that are highly adaptable, wherein decisions are taken in a decentralized way. Cooperation also favors the creation of products designed for and adapted to new needs and demands, and the development of a more efficient process of innovation that incorporates the "voice of the consumer" together with the experience and know-how of other partners. It also reduces the uncertainty surrounding the product's future and its dependence at the time of product launch, while improving on the results obtained.

Yet, it has been estimated that approximately 60% of established cooperative relationships fail, and there are various factors that have been identified as barriers to effective cooperation. Lack of familiarity between the partners, the distance that separates them, and the absence of prior collaboration experience are noted as the most important inhibitors of the process of cooperation. The issue of lack of familiarity arises when the primary organization fails to research what each partner's desired benefits, level of risk aversion, level of commitment, and strategic similarity are prior to the beginning of the project. The issue of distance can be physical, time related, or cultural. The last inhibitor, the absence of prior collaboration experience, arises when partners have not been a part of these types of alliance previously. The thought is that partners with prior experience will be more able to efficiently and effectively partner with organizations in new alliances making the overall management of the relationship easier for all parties (Vilaseca-Requena, Torrent-Sellens, and Jimenez-Zarco, 2007).

Partner selection and management is inherently important. Prior to beginning a project, each party should agree to the specific benefits to be gained from the relationship, as well as the risks and compromises they are willing to accept. When selecting partners to join these types of alliances, special consideration should be given to whether the partner has the necessary resources/capacity to meet agreed commitments, whether the partners' culture or strategy is compatible with the primary organization, and whether the help of the partner can increase the efficiency and efficacy of the innovation process (Vilaseca-Requena, Torrent-Sellens, and Jimenez-Zarco, 2007).

Product platforms. Expanding internationally can be a difficult and costly task. To circumvent some of the costs and problems associated with this, many firms use product platforms. Broadly defined, *a platform is a set of product components that are physically connected as a stable subassembly and are common to different final models* (Muffatto, 1999). In other words, it is a foundation that can be used to create several different final products. The automobile industry uses product platforms for several components that are used in a variety of their models. The product platform concept represents a powerful approach for manufacturers to compete cost-effectively in a global market that requires diverse product range, quick time to market, and rapid responses to supply sources (Zhang, 2008).

There are many benefits to using product platforms. It creates flexibility by allowing companies to produce multiple product variations with limited impact on production and assembly processes. It also reduces the need for a large number of parts, which in turn reduces the amount of suppliers needed, and reduces costs. Another key benefit derived from the use of product platforms is a reduction in lead-time, and the reduction or elimination of many preassembly operations that reduces the throughput time. This helps companies react to market changes faster.

The biggest challenge in using product platforms is how to strike a balance between commonality and modularity. In other words, how common can a product line be while still creating enough variations to satisfy the global

market? The concept of platforming enables the manufacturer to further "commonize" the product family into fewer variants in order to take advantage of economies of scale (Zhang, 2008). However, platform approaches often result in the reduction of the range of customer choices, which can hinder overall sales.

Global product development teams. There is further complexity involved when broad geographical considerations are added to the innovation management and NPD equation. Global new product development teams are often established to address the needs of common global markets, to incorporate the unique needs of local markets, and to bring together globally diverse resources and expertise (Barczak and McDonough, 2003). Yet, these teams are often difficult to manage because of geographic and cultural diversity; these groups can achieve a higher level of performance if there is a significant degree of information exchange (Cummings, 2004). There can also be conflict between functional group members such as engineering and marketing (Maltz and Kohli, 2000), which impedes the effective development and ultimate success of new products – this effect can be exacerbated in culturally diverse groups. Thus, it is argued that the way work centers are structured and their relationship to the international network should be based on the underlying characteristics of a firm's knowledge-based assets (Birkinshaw, 2002) and the global strategic orientation of the firm.

Conclusion

Managing global product innovation and development presents quandaries for international companies with respect to balancing the need for within-country representation with between-country comparability. As efficient production processes themselves became commonplace, product differentiation emerged as a key in creating a sustainable competitive advantage. This article reviews some of the more salient issues faced by managers of product innovation and development in a global marketplace. Global strategic orientation of the firm plays a key role in determining the nature of product development, interacting with the degree of global market segmentation for the industry to determine the relative level of product standardization or adaption to be supported. This is a key point because of the costs and benefits that can accrue to the firm through innovation. Organizational design structures and elements support the NDP processes, and are particularly complex for global products and variations. As the world continues to become more integrated through technological advancement, governmental initiatives, and infrastructure improvements, global product innovation and development will evolve to meet the challenges of the marketplace.

Acknowledgment

Special thanks to Angel Lynch and Shane Meldrum for their help on an earlier version of this article, and to Ahmet Kirca for critical commentary.

Bibliography

Arranze, N. and Arroyabe, J.C. (2008) The choice of partners in R&D cooperation: an empirical analysis of Spanish firms. *Technovation*, **28**, 88–100.

Barczak, G. and McDonough, E.F.III (2003) Leading global product development teams. *Research Technology Management*, **46** (6), 14–18.

Bartlett, C.A. and Ghoshal, S. (1987) Managing across borders: new strategic requirements. *Sloan Management Review*, **28** (4), 7–17.

Birkinshaw, J. (1996) How multinational subsidiary mandates are gained or lost. *Journal of International Business Studies*, **27** (3), 467.

Birkinshaw, J. (2002) Managing internal R&D networks in global firms – what sort of knowledge is involved? *Long Range Planning*, **35** (3), 245.

Birkinshaw, J., Morrison, A., and Hulland, J. (1995) Structural and competitive determinants of a global integration strategy. *Strategic Management Journal*, **16** (8), 637–655.

Boudreau, M.-C., Loch, K.D., Robey, D., and Straud, D. (1998) Going global: using information technology to advance the competitiveness of the virtual transnational organization. *Academy of Management Executive*, **12** (4), 120–128.

Buzzell, R. (1968) Can you standardize multinational marketing. *Harvard Business Review*, , **46** 103–113.

Calantone, R.J., Cavusgil, S.T., and Schmidt, J.B. (2004) Internationalization and the dynamics of product adaptation: an empirical investigation. *Journal of Product Innovation Management*, **21**, 185–198.

Calantone, R.J. and DiBenedetto, A.C. (1995) Business performance and strategic new product development activities: an empirical investigation. *Journal of Product Innovation Management*, **12**, 214–223.

Cateora, P. and Graham, J. (2009) *International Marketing*, 14th edn, McGraw Hill.

Cavusgil, S.T., Zou, S., and Naidu, G.M. (1993) Product and promotion adaptation in export ventures: an empirical investigation. *Journal of International Business Studies*, **24** (3), 479–506.

Clegg, A. (2008) Market Research: An Open Door Policy. *Marketing Week* (Jan 3), p. 17.

Cooper, E.K. (1991) The impact of product innovativeness on performance. *Journal of Product Innovation Management*, **8**, 240–251.

Cummings, J.N. (2004) Work groups, structural diversity, and knowledge sharing in a global organization. *Management Science*, **50** (3), 352–365.

Dawar, N. and Parker, P. (1994) Marketing universals: consumers' use of brand name, price, physical appearance, and retailer reputation as signals of product quality. *Journal of Marketing*, **58**, 81–95.

Frost, T.S., Birkinshaw, J.M., and Ensign, P.C. (2002) Centers of excellence in multinational corporations. *Strategic Management Journal*, **23** (11), 997.

Ghoshal, S. (1987) Global strategy: an organizing framework. *Strategic Management Journal*, **8** (5), 425–440.

Harmancioglu, N., Calantone, R.J., McNally, R., and Durmusoglu, S. (2007) Your new product development (NPD) is only as good as your process: an exploratory analysis of new NPD process design and implementation. *R & D Management*, **37**, 399–424.

Jain, S.C. (1989a) *Export Strategy*, Quorum Books, New York.

Jain, S.C. (1989b) Standardization of international marketing strategy: some research hypothesis. *Journal of Marketing*, **53** (1), 70–79

Johanson, J. and Vahlne, J.-E. (1977) The internationalization process of the firm: a model of knowledge development and increasing foreign market commitments. *Journal of International Business Studies*, **8**, 23–32.

Johansson, J.K. and Yip, G.S. (1994) Exploiting globalization potential: U.S. and Japanese strategies. *Strategic Management Journal*, **15** (8), 579–601.

Katz, B. (1987) *Managing Export Marketing*, Aldershopt, Hants, England.

Kuvykaite, M.A. (2008) Standardization/adaptation of marketing solutions in companies operating in foreign markets: an integrated approach. *Engineering Economics*, **1** (56), 37–47.

Majaro, S. (1982) *International Marketing: A Strategic Approach to World Markets*, Allen and Unwin, London, England.

Maltz, E. and Kohli, A. (2000) Reducing marketing's conflict with other functions: the differential effects of integrating mechanisms. *Journal of the Academy of Marketing Science*, **28** (4), 479–492.

Martinez, J.I. and Jarillo, J.C. (1991) Coordination demands of international strategies. *Journal of International Business Studies*, **22** (3), 429.

Muffatto, M. (1999) Platform strategies in international new product development. *International Journal of Operations and Production Management*, **19**, 449–459.

Ohmae, K. (1989) Managing in a borderless world. *Harvard Business Review*, **67**, 152–161.

Perlmutter, H.V. (1969) The tortuous evolution of the multinational corporation. *Columbia Journal of World Business*, **4**, 9–18.

Pimblett, J.W. (1997) The standardization debate in marketing strategies: a construct and a research agenda. *Journal of the Academy of Marketing Science*, **19** (1), 1–10.

Roth, K., Schweiger, D.M., and Morrison, A.J. (1991) Global strategy implementation at the business unit level: operational capabilities and administrative mechanisms. *Journal of International Business Studies*, **22** (3rd Quarter), 369–402.

Rothwell, R. (1994) Towards the fifth-generation innovation process. *International Marketing Review*, **11**, 7–31.

Townsend, J.D., Calantone, R.J., and Schmidt, J.B. (2003) *Foreign Impact: A Longitudinal Study of the Liability of Foreignness in the U.S. Motion Picture Market*, American Marketing Association Summer Educators Conference, Chicago.

Townsend, J.D., Yeniyurt, S., Deligonul, S., and Tamer Cavusgil, S. (2004) Marketing related antecedents of performance in the global company. *Journal of International Marketing*, **12** (4), 1–24.

Townsend, J.D., Yeniyurt, S., and Talay, M.B. (2009) Getting to global: an evolutionary perspective of brand expansion in international markets. *Journal of International Business Studies*, **40** (2), 301–320.

Usunier, J.C. (2003) *Marketing Across Cultures*, Financial Times Prentice Hall, London.

Vilaseca-Requena, J., Torrent-Sellens, J., and Jimenez-Zarco, A.I. (2007) ICT use in marketing in innovation success factor: enhancing cooperation in new product development processes. *European Journal of Innovation Management*, **10** (2), 268–288.

Wind, Y. (1986) The myth of globalization. *Journal of Consumer Marketing*, **3** (2), 23–26.

Yavas, U., Verhage, B.J., and Green, R.T. (1992) Global consumer segmentation versus local market orientation: empirical findings. *Management International Review*, **32**, 265–272.

Yeniyurt, S. and Townsend, J.D. (2003) Does culture explain acceptance of new products in a country? An empirical investigation. *International Marketing Review*, **20** (4), 377–396.

Yip, G.S. (1992) *Total Global Strategy: Managing for Worldwide Competitive Advantage*, Prentice Hall, New Jersey.

Zhang, X. (2008) Simultaneous configuration of platform products and manufacturing supply chains. *International Journal of Production Research*, **46** (6), 137–162.

Zou, S. and Cavusgil, S.T. (2002) The GMS: a broad conceptualization of global marketing strategy and its effects on firm performance. *Journal of Marketing*, **66** (4), 40–56.

managing the global product portfolio

Warren J. Keegan and Elyse Arnow Brill

Introduction

Managing the global product portfolio has been a topic of discussion and debate in the marketing literature for the past five decades. Commencing with global advertising (*see* INTERNATIONAL ADVERTISING – IS THERE STILL A STANDARDIZATION VERSUS LOCAL ADAPTATION DEBATE?) in the early 1960s, the entire MARKETING MIX has been included in the standardization versus adaptation debate. There is, however, among proponents of either approach, the fundamental recognition that the discipline of marketing is universal, that is, it is a set of concepts, tools, theories, and practices that together constitute a body of knowledge that can be applied to varying competitive markets (*see* COMPETITIVE ANALYSIS) around the world.

The Strategic Concept of Marketing

While the marketing discipline is universal and global, markets are differentiated. Market conditions vary from world region to region, from country to country, and from segment to segment within countries. Marketing objectives and target-market scope may be directed to local, national, regional, or global areas. Indeed, there are today BORN GLOBAL firms that have never had a primary, domestic market (Knight and Cavusgil, 2004).

What makes this debate possible is the evolution of the marketing concept from one focused on the sale of a "better" product defined by standards and values determined by the company, to a strategic concept of marketing that has shifted the focus from the product and its interaction with the consumer to everything. The global marketer must know the customer in a context defined by existing and future competition (*see* COMPETITIVE ADVANTAGE: ITS SOURCES AND THE SEARCH FOR VALUE), regulation, government policy, and the influences of economic, technical, social (*see* SOCIAL CLASS; CHILDHOOD SOCIALIZATION AND INTERGENERATIONAL INFLUENCES; SOCIAL INFLUENCE; SOCIAL NETWORKS), cultural (*see* SUBCULTURES; CROSS-CULTURAL PSYCHOLOGY OF CONSUMER BEHAVIOR; MARKETING ASPECTS OF CULTURAL DISTANCE; BASE OF THE PYRAMID MARKETS: CULTURE INSIGHTS AND MARKETING IMPLICATIONS), and political macro forces that shape markets and their evolution. Everything matters. In the global auto industry, for example, successful competitors must meet world standards of functionality and performance, as well as world class appeal of design, fit, and finish that appeal to customer emotions. This means that the successful global competitor must have designers, stylists, and engineers who communicate with each other and who are aware of market demands in every world region that the company addresses in its global marketing effort. The successful competitor must also be price competitive (*see* PRICING STRATEGY; INTERNATIONAL PRICING OBJECTIVES AND STRATEGIES), and this requires the elimination of unnecessary differentiation and duplication of effort.

Integrated Global Marketing

An important issue in global strategic marketing is whether theory and experience based in national markets is applicable to global marketing (Hongsik, Change-Hoan, and Sutherland, 2007). Discussion has focused on the benefits and/or drawbacks of standardization (*see* BASE OF THE PYRAMID MARKETS: CULTURE INSIGHTS AND MARKETING IMPLICATIONS), arguments for or against a growing homogeneity of markets and customer needs, contextual factors having an impact on the degree of MARKETING MIX, standardization, and performance outcomes of the standardization and adaptation approaches (Birnik and Bowman, 2007).

Howard Perlmutter identified three managerial orientations in his seminal 1969 article, The Tortuous Evolution of the Multinational Corporation (Perlmutter, 1969). These orientations included an ethnocentric or home-country orientation, a polycentric or host-country orientation, and a geocentric or world orientation. Almost twenty years later, Bartlett and Ghoshal's (Bartlett and Ghoshal, 1989) typology of organizational development (domestic, international, multinational, and

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

global) reflect different orientations and strategies of globalizing companies. With reference to the marketing function, Sheth, in 2001, and others before him, postulated that in light of important macroeconomic forces shaping the world and its markets, a "borderless world" was emerging where national boundaries were becoming less relevant to marketing and marketing planning practices (Sheth, 2001a). Sheth suggested that *intra*national differences were overtaking *inter*national differences; for the global marketer, *within-country* differences were displacing *between-country* differences in importance. Accordingly, he suggested strategic marketers consider different target markets within a country, similar to those that might exist on a global basis.

By focusing in this manner, a company's perspective shifts from between-country differences to transnational similarities across national boundaries. In Sheth's words, "the elitism of the diffusion in innovations from the home country to the rest of the world [is] replaced with the concept of worldwide markets" (*see* DIFFUSION OF INNOVATION); (Sheth, 2001b). MARKETING PLANNING would provide target-market coverage on a worldwide basis (Sheth, 2001c).

A second shift at the organizational level supports this focus on transnational similarities with the goal of creating value (*see* CUSTOMER SOLUTIONS; A FRAMEWORK FOR CREATING VALUE PROPOSITIONS) for targeted customers. This dual geocentric view at the organizational and customer levels demands, among other things, continued efforts to lower costs and increase learning and knowledge sharing across the global enterprise.

The Standardization/Adaptation Debate

Buzzell offered one of the first theoretical discussions of standardization as an international marketing strategy in 1968 (Buzzell, 1968). He argued that dissimilarities among countries led multinational firms to perceive and design MARKETING PLANNING in each country as a local problem with localized solutions. Buzzell provided several rationales for an international standardization strategy (*see* STANDARDIZATION/ADAPTATION OF INTERNATIONAL MARKETING STRATEGY), most importantly, cost savings arising from standardization of the MARKETING MIX elements and the creation of a single marketing strategy spanning the globe. He applied standardization strategy to the marketing mix elements – the 4 Ps of product, pricing, promotion, and place, highlighting potential economies in PRODUCT DESIGN decisions, (*see* GLOBAL PRODUCT DEVELOPMENT), advertising (*see* INTERNATIONAL ADVERTISING – IS THERE STILL A STANDARDIZATION VERSUS LOCAL ADAPTATION DEBATE?), packaging and messaging (*see* INTEGRATED MARKETING COMMUNICATION STRATEGY) relating to brand development (*see* BRAND GROWTH STRATEGY; BRAND STRATEGY; BRAND VALUE) as well as organizational/operational elements including stronger planning, control, and the use of ideas with transnational appeal in as many markets as possible.

Among the many obstacles to standardization, Buzzell noted important market characteristics including a target market's stage of economic and industrial development, the myriad of existing attitudinal (*see* ATTITUDES) and cultural impediments (*see* CROSS-CULTURAL PSYCHOLOGY OF CONSUMER BEHAVIOR; SUBCULTURES) as well as product-use conditions and differences in marketing infrastructure (*see* MARKET/INDUSTRY STRUCTURE) including channels of distribution (*see* MARKETING CHANNEL STRATEGY; DESIGNING A GLOBAL SUPPLY CHAIN: OPPORTUNITIES AND CHALLENGES) and media (*see* INTEGRATED MARKETING COMMUNICATION STRATEGY). Buzzell concluded that the pros and cons of standardization should be considered when designing and implementing a global marketing strategy. Decisions to standardize, he recommended, should be based on the net impact of standardization on estimated overall revenue and projected costs. "Obviously, each case must be considered on its own merits. . . ." (Buzzell, 1968).

Buzzell's original question with respect to global strategy continues to ring true today. "Is it practical to consider the development of a marketing strategy in terms of all of its

elements on an integrated, global marketing scale?" (Buzzell, 1968).

As noted just over 30 years later, "customization" and "fine-tuning" of international marketing are included in the concept of global marketing (Sheth and Parvatiyar, 2001). Its underlying philosophy is standardization of marketing programs wherever possible and customization wherever necessary (Sheth and Parvatiyar, 2001). Keegan, in an article titled Multinational Product Planning: Strategic Alternatives published in 1969, identified four strategic alternatives using the marketing mix based on product-standardization versus adaptation and promotion-standardization versus adaptation dimensions (Keegan 1969). He concluded that the "choice of communications strategy . . . is a function of three key factors: 1. the product itself defined in terms of the function or need it serves, 2. the market defined in terms of the conditions under which the product is used . . . and, 3. the costs of adaptation and manufacture." (*see* PRODUCT DESIGN; MARKET DEFINITION) Much literature in the field of international marketing has built on this multidimensional perspective (Quelch and Hoff, 1986) and continues to look beyond single strategy dimensions.

Theodore Levitt's seminal 1983 article, "The Globalization of Markets," has often been characterized as arguing that in the face of diminishing cultural (*see* MARKETING ASPECTS OF CULTURAL DISTANCE; BASE OF THE PYRAMID MARKETS: CULTURE INSIGHTS AND MARKETING IMPLICATIONS) and social differences (*see* SOCIAL CLASS; SOCIAL INFLUENCE) across countries, a standardized marketing strategy will increase profitability and competitiveness (Levitt, 1983). Levitt did assume that demand was becoming more homogenous worldwide; global consumers (*see* SOCIETY, CULTURE, AND GLOBAL CONSUMER CULTURE; GLOBAL CONSUMERISM AND CONSUMPTION), he pointed out, were increasingly demanding high-quality, low-cost products and services. Driving this trend, he argued, were global communications, travel, and the emergence of a global consumer with shared experience, knowledge, and expectations. Although Levitt forcefully argued for the recognition of this trend, he did not ever call for standardization per se. Like Buzzell and Keegan, he recognized that marketers should standardize where they could, and adapt where it was needed. He noted, for example,

> Large companies operating in a single nation or even a single city don't standardize everything they make, sell or do, [t]hey have product lines instead of a single product version, and multiple distribution channels. There are neighborhood, local, regional, ethnic, and institutional differences even within metropolitan areas. . . . Success in a world with homogenized demand requires a search for sales opportunities in similar segments across the globe . . . [T]he most effective world competitors incorporate superior quality and reliability into their cost structures . . . They compete on the basis of appropriate value – the best combinations of price, quality, reliability, and delivery for products that are globally identical with respect to design, function and even fashion (Levitt, 1983) (*see* MARKET SEGMENTATION AND TARGETING; VALUE PROPOSITION; CUSTOMER SOLUTIONS.

Levitt acknowledged the creation of customer value (*see* CUSTOMER SOLUTIONS) as the basis of competition. Levitt's point was not that standardization should be the uniform practice of global marketers. Instead, he underlined the importance of creating value (*see* VALUE PROPOSITION) by offering functional utility, quality, design, and price.

Proponents of a strategy of adaptation for globalizing firms highlight the obvious dissimilarities between and among markets of differing countries and regions and urge the global marketer to differentiate marketing programs. Such market diversity includes a full set of macro environmental factors including geographical, topographical and climate constraints, differing political and economic influences, and varying important cultural imperatives (*see* CROSS-CULTURAL PSYCHOLOGY OF CONSUMER BEHAVIOR) reflected in language, laws and regulation, and religious beliefs along with race and ethnicity considerations critical to successful marketing programs (Czinkota and Ronkainen, 1998). Proponents of this view urge that adjusting the determinants of the

MARKETING MIX and marketing strategy (*see* MARKETING STRATEGY IMPLEMENTATION) are imperative to comply with local custom, culture, and taste. Adapting market offerings to different markets, proponents argue, addresses the "friction" that necessarily exists between differing stakeholders because of environmental differences, thereby permitting better overall performance (Shoham, 2002).

The debate between the two "views" has been a tempest in a teapot. Levitt did not ignore the real differences within and between global markets. He identified the changes in globalizing markets that had created an opportunity to provide value (*see* VALUE PROPOSITION) to global customers (*see* SOCIETY, CULTURE, AND GLOBAL CONSUMER CULTURE) by not only responding to differences but also by recognizing similarities and opportunities to reduce costs and increase quality through greater design unification. Success for globalizing international firms requires both adapting to unique differences and standardizing elements of the marketing mix that do not need to be adapted to meet customer needs. These marketing planning decisions highlight the trade-offs in each market between economies associated with standardization and the value added by differentiation (Viswanathan and Dickson, 2007).

The New Paradigm: Multidimensional Global Marketing Strategy

Successful globalizing firms have adopted a strategy of integrating standardization, adaptation (*see* STANDARDIZATION/ADAPTATION OF INTERNATIONAL MARKETING STRATEGY), and innovation in their marketing programs. For every component of marketing strategy and program planning, analysis is required to understand the optimal standardization/adaptation level. The recent development of both contingency and configurational theories have expanded the binary standardization/adaptation analysis and brought into focus the important contextual nature of global marketing strategy (Nanda, Viswanathan, and Dickson 2007).

Contingency theory posits that no universal set of strategic choices is most suited for all circumstances or organizations; strategic choices are situation specific. Configurational theory holds that structural and program effectiveness arises from the superior combinations of strategic and structural elements, which, when combined in various ways, interact differently. These superior combinations have predictive capabilities for organizational behavior and success (Lim, Acito, and Rusetski, 2006).

Recent literature suggests variables influencing strategic marketing choices. Internal organizational elements include, for example, the degree of centralization of decision making and the international experience of the firm. Strategic variables highlight market entry modes (*see* MARKET ENTRY AND EXPANSION) and business goals and objectives. External elements include the size of the local market, degree of local competitive intensity and similarity of markets including their stage of economic development (*see* STAGES OF MARKET DEVELOPMENT), "psychic" or cultural distance (*see* MARKETING ASPECTS OF PSYCHIC DISTANCE; MARKETING ASPECTS OF CULTURAL DISTANCE) between markets, and similarity of legal and regulatory constraints and marketing infrastructure (Birnik and Bowman, 2007). Various classifications addressing multidimensional analysis have been highlighted over the years. In 2009, Nasir and Altinbasak proposed a conceptual framework encompassing many of these external and internal drivers (Nasir and Altinbasak, 2009).

The importance of identified contingency factors, whether internal or external, will vary with each MARKETING PLANNING decision. In light of this fact, differing elements of the MARKETING MIX reflect various levels of standardization, adaptation, or innovation. For example, there is evidence that the optimum level of standardization of the marketing process is higher than the standardization of marketing program content (Birnik and Bowman, 2007). Research into pricing patterns (*see* PRICING STRATEGY) shows this to be the least or one of the least standardized elements of the marketing mix with sales, distribution, and promotion not far behind. Studies indicate that packaging patterns reflect medium-to-high levels of standardization. Similar medium levels of standardization are applicable to advertising although research distinctions have been made

at the level of the basic advertising message, the creative message, and advertising content material as compared to preplanned design efforts.

The product is the element of the marketing mix with the highest potential for standardization. Industrial products typically offer more potential for standardization than consumer products because they are less influenced by local customs, taste, and habits. Major consumer durables offer greater potential for standardization because the components of the product are not visible to the consumer, and perform a function that is not influenced or affected by cultural differences. The automobile is a good example of a consumer durable with high standardization potential. All of the successful or would-be-successful global or globalizing automobile companies recognize this fundamental fact of their industry.

Conclusion

While the marketing discipline is universal, the heterogeneity of subnational, national, and regional markets does not allow complete standardization of international marketing planning or programming. It is necessary to adapt to certain differences in markets. On the other hand, similarities in markets can and, indeed, must be addressed with standardized elements of the marketing offering. The task of the global marketer is to adapt where adaption is needed and is worth the expense, and where it strengthens competitive advantage, and to standardize where it is not needed and adds unnecessary cost.

A multidimensional approach embraces important contextual, historical, and competitive factors found in diverse markets. Taking this approach, global marketing strategy plays an increasingly important role in determining both the strategic direction of a firm and its worldwide growth success or failure. The ultimate test of success is the creation of unique value where value is defined by the sweet spot where the benefits and price of the market offering afford the greatest value to the customer. Successful global companies have developed the ability to find that sweet spot, and have been rewarded by competitive success in global markets.

Bibliography

Albaum, G. and Tse, D. (2001) Adaptation of international marketing strategy components, competitive advantage and firm performance: a study of Hong Kong exporters. *Journal of International Marketing*, **9** (4), 59–81.

Bartlett, C. and Ghoshal, S. (1989) *Managing Across Borders: The Transnational Solution*, Harvard Business School Press, Boston.

Birnik, A. and Bowman, C. (2007) Marketing mix standardization in multinational corporations: a review of the evidence. *International Journal of Management Reviews*, **9** (4), 303–324. For a "summary model" of the contextual variables found to differentiate strategies of higher marketing mix standardization from those with lower standardization.

Boddewwyn, J.J., Soehl, R., and Picard, J. (1986) Standardization in international marketing: is ted levitt in fact right?. *Business Horizons*, **29** (6), 61–79.

(a) Buzzell, R.D. (1968) Can you standardize multinational marketing? *Harvard Business Review*, **46** (6), 102–133; (b) Buzzell, R.D. (1968) Can you standardize multinational marketing? *Harvard Business Review*, **46** (6), 10; (c) Buzzell, R.D. (1968) Can you standardize multinational marketing? *Harvard Business Review*, **46** (6), 1.

Cateora, P.R. and Graham, J.L. (1999) *International Marketing*, 10th edn, Irwin McGraw-Hill, Boston.

Cho, C.H. and Cheon, H.J. (2005) Cross-cultural comparisons of interactivity on corporate web sites. *Journal of Advertising*, **34** (2), 99–115.

Christensen, C. (1998) *Innovator's Dilemma*, Harvard Business School, Boston.

Czinkota, M.R. and Ronkainen, I.A. (1998) *International Marketing*, 5th edn, The Dryden Press, Fort Worth, London.

Hongsik, J.C., Change-Hoan, C., and Sutherland, J. (2007) A meta-analysis of studies on the determinants of standardization and localization of international marketing and advertising strategies. *Journal of International Consumer Marketing*, **19** (4), 109.

Keegan, W.J. (1969) Multinational product planning: strategic alternatives. *Journal of Marketing*, **19** (1), 58–62.

Keegan, W.J., (2002), *Global Marketing Management*, 7th edn, International Edition, Prentice Hall, New Jersey.

Knight, G.A. and Cavusgil, S.T. (2004) Innovation, organizational capabilities and the born-global firm. *Journal of International Business Studies*, **35** (2), 124–141.

Levitt, T. (1983) The globalization of markets. *Harvard Business Review*, **61** (3), 92–102.

Lewis, K.S., Lim, F.A., and Rusetski, A. (2006) Citing Miller and Mintzberg (1983), The case for configuration, in *Beyond Method: Strategies for Social Research* (ed. G., Morgon), Sage, Beverly Hills, pp. 57–73.

Lim, L., Acito, F., and Rusetski, A. (2006) citing Miller and Mintzberg (1983), the case for configuration, in *Beyond Method: Strategies for Social Research* (ed. G. Morgon), Sage, Beverly Hills, pp. 57–73.

Nanda, K., Viswanathan, P., and Dickson, R. (2007) The fundamentals of standardizing global marketing strategy. *International Marketing Review (London)*, **24** (1), 46. See, for example, discussion of an expanded set of factors including the level of homogeneity of customer response to the marketing mix, the transferability of competitive advantage, and the homogeneity of economic freedom across markets.

Nasir, A. and Altinbasak, I. (2009) The standardization/adaptation debate: creating a framework for the new millennium. *Strategic Management Review*, **3** (1), 17.

Ohmae, K. (1990) *The Borderless World*, HarperBusiness, New York.

Perlmutter, H. (1969) The tortuous evolution of the multinational corporation. *Columbia Journal of World Business*, **4**, 9–18.

Quelch, J.A. and Hoff, E.J. (1986) Customizing global marketing. *Harvard Business Review*, **64** (3), 59–68.

(a) Sheth, J.N. (2001a) From international to integrated marketing. *Journal of Business Research*, **51** (1), 5–9; (b) Ohmae, K. (1990) *The Borderless World*, Harper Business, New York.

Sheth J.N. (2001b) From international to integrated marketing. *Journal of Business Research*, **51** (1), 5–9.

(a) Sheth, J.N. (2001c) From international to integrated marketing. *Journal of Business Research*, **51** (1), 5–9; (b) Christensen, C. (1998) *Innovator's Dilemma*, Harvard Business School, Boston.

Sheth, J.N. and Parvatiyar, A. (2001) The antecedents and consequences of integrated global marketing. *International Marketing Review (London)*, **18** (1), 16.

Shoham, A. (2002) Standardization of international strategy and export performance: a meta-analysis. *Journal of Global Marketing*, **16** (1/2), 97–120.

Viswanathan, N. and Dickson, P. (2007) The fundamentals of standardizing global marketing strategy. *International Marketing Review (London)*, **24** (1), 46.

Vrontis, D. and Thrassou, A. (2007) Adaptation vs. standardization in international marketing–the country of origin effect. *Innovative Marketing*, **3** (4), 7.

global branding: three keys for global brand success

Kevin Lane Keller

Introduction

Many companies have been global marketers for decades – firms like Nestlé, Shell, Bayer, and Toshiba have sold their products around the world for years. In more and more product categories, the ability to establish a global profile is becoming virtually a prerequisite for success. In luxury goods such as jewelry, watches, and handbags, where the addressable market is a relatively small percentage of the global market, a global profile is essential to grow profitably. Marketers for luxury brands such as Prada, Gucci, Cartier, and Louis Vuitton have long managed lucrative global franchises.

Besides the need for a larger customer base to achieve necessary economies of scale, companies may look to sell outside their domestic market for a number of different reasons: better perceived profit opportunities in international markets than in the domestic market; a need to diversify risk and reduce their dependence on any one market; a desire to counterattack global competitors in their home markets; and a realization that their customers are going abroad and require international service.

But global competition is intensifying as new firms make their mark on the international stage. The automotive market is becoming a worldwide free-for-all. In Chile, for example, because there are no domestic auto manufacturers, imports come from all over the world, including 14 different brands of Chinese cars, trucks, and commercial vehicles. Competition in developing markets has also intensified. In China's exploding mobile-phone market, Motorola found their market share drop in half over a two-year period because of inroads made by Nokia and different Asian competitors.

Competition arising from firms based in developing markets is also heating up (*see* EMERGING MARKETS). In various developing markets, India's Tata Motors have launched the people's car whose spartan features are offset by a rock-bottom price. Eyeing more developed markets down the road, Tata can afford to charge a fraction of what other auto manufacturers charge because of their reduced development costs and innovative distribution strategy that requires dealers to participate in the final assembly.

India's Mahindra Motors are not even going to wait before entering developed markets. Their four-door, diesel-powered short-bed trucks are tackling the competitive small utility vehicle (SUV) and truck markets in Europe, Asia, and the United States with a promise of superior fuel economy. To offset a lack of image and reputation, Mahindra are targeting three groups in the United States that are believed to be most receptive to their appeals: consumers who identify themselves as "green"; people who have bought their other main automotive product, Mahindra tractors; and Indian expatriates.

Given the significant growth opportunities offered by international markets, developing a global strategy can be of paramount importance to brand builders everywhere. For many companies, however, global branding has been a mixed blessing. On the one hand, a global branding program can lower marketing costs, realize greater economies of scale in production, increase distribution efficiencies, and provide a long-term source of revenue. On the other hand, if not designed and implemented properly, a global branding program may fail as a result of ignoring important differences in consumer behavior and/or the competitive environment in the individual countries.

The goal for any brand builder, obviously, is to obtain as many of the benefits of global branding as possible while minimizing the potential risks and downside. Not surprisingly, many companies have experienced both tremendous success and embarrassing failures in their global branding efforts. It is not always the case that the most successful brand in one country will find success in other countries. Although US retail giant Wal-Mart have had some success entering the overseas markets in Latin America and China, despite concerted efforts, they found themselves having to withdraw from both the German and South Korean markets.

The goal of this article is to share some common themes or guidelines for success that have emerged in global branding (see Johansson, 2009 for more detail). We outline three keys

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

for global brand success: (i) understand the global consumer context, (ii) build a solid global marketing foundation, and (iii) strike a balance in global brand management. They represent global branding fundamentals that provide the necessary foundation for creating a strong global brand. We raise a number of different issues and offer a number of different examples in each case.

Understand the Global Consumer Context

First – and perhaps most fundamental – it is important to recognize that international markets vary in terms of consumer behavior and all the different marketing forces and other factors that impact them (*see* BASE OF THE PYRAMID MARKETS: CULTURE INSIGHTS AND MARKETING IMPLICATIONS). As a result, consumers may vary accordingly in their perceptions, beliefs, attitudes, images, experiences, behaviors, and so on, toward both the brand itself as well as the product category in general. These differences can have profound implications on building and managing brand equity across geographical boundaries.

For example, consider the following (Hollis, 2008). The median age in India and China is roughly 25 years, whereas in Japan, Germany, and Italy it is around 43. When asked if they are more concerned with getting a specific brand than the best price, roughly two-thirds of Americans agreed as compared to around 80% in Russia and India. A lot of these differences in consumer behavior reflect cultural differences that can be pronounced across countries. Hofstede (1980) identifies four cultural dimensions that differentiate countries (with countries or areas that score high and low):

1. *Individualism versus collectivism*. In collectivist societies, the self-worth of an individual is rooted more in the social system than in individual achievement (high: Japan; low: United States).
2. *High versus low power distance*. High power distance cultures tend to be less egalitarian (high: Russia; low: Nordic).
3. *Masculine versus feminine*. This dimension measures how much the culture is dominated by assertive males versus nurturing females (high: Japan; low: Nordic).
4. *Weak versus strong uncertainty avoidance*. Uncertainty avoidance indicates how risk aversive people are (high: Greece; low: Jamaica).

At the same time, many countries do not vary much on one or more of these various considerations, suggesting that differences in marketing activity can create unnecessary or ineffective marketing activity. An important key to global marketing success is understanding consumers in different markets, recognizing what they know and feel, and could potentially value about the brand, and, as described below, tailoring marketing programs to their desires as a result.

Obviously, the fewer the differences in consumer behavior found across markets, the more effective a standardized marketing program will be. Some types of products travel better across borders than others. New products often represent promising candidates for standardization. While mature products may have vastly different histories (or even positions) in different markets, consumer knowledge for new products is generally the same everywhere because perceptions are yet to be formed. Many of the leading internet brands – Google, eBay, and Amazon – have made relatively quick progress in overseas markets. In addition, high-end products also benefit from standardization because a high quality or prestige image often can be marketed similarly across countries. On the other hand, food and beverage marketers find it more challenging to standardize their products as they have to contend with widely varying tastes and cultural habits and norms.

Developing markets. In understanding consumer behavior in a global context, because of the wide income and economic disparity involved, some of the biggest differences are found between consumers in developing or emerging (e.g., the BRICS countries: Brazil, Russia, India, China, and South Africa) versus developed markets (Mahajan, De Moraes, and Wind, 2000; Khanna and Palepu, 2006). These differences can have profound implications forhow brands should be marketed (*see* EMERGING MARKETS).

For example, consider channels of distribution. Eighty percent of consumers in developing markets buy their products from tiny bodegas, stalls, kiosks, and mom-and-pop stores not much bigger than a closet, which Procter & Gamble call *high-frequency stores*. Smaller packaging and lower sales prices are often critical when incomes and housing spaces are limited. Unilever's 4-cent sachets of detergent and shampoo have been successful in rural India, where 70% of the country's population still lives. Coca-Cola's sales jumped when they moved to a smaller, 200-ml bottle in India, selling it for 10–12 cents in small shops, bus-stop stalls, and roadside eateries. Recognizing that their cost structure made it difficult to compete effectively in developing markets, Procter & Gamble have devised a number of cheaper, clever ways to make the right kinds of products to suit consumer demand there.

Fundamentally, marketers must rethink all aspects of their marketing program in developing markets. As another example, high cell-phone penetration in developing markets makes mobile marketing an attractive option. A pioneer in China, Coca-Cola China created a national campaign asking Beijing residents to send text messages guessing the high temperature in the city every day for just over a month, for a chance to win a one-year supply of Coke products. The campaign attracted more than four million messages over the course of 35 days. In Africa, mobile-phone operator Celtel invested in rural services by introducing the Me2U service, by which callers could send airtime credit to other mobile phones. Because most Africans do not have bank accounts, it has become a convenient and cheap way to transfer money, even substituting for cash in some villages.

Build a Solid Global Marketing Foundation

The second guideline emphasizes the importance of building a solid global marketing foundation. A solid global marketing foundation results when (i) the proper marketing infrastructure is put into place; (ii) the right marketing partners are enlisted; and (iii) steps are not skipped in brand building.

Proper marketing infrastructure. A critical success factor for many brands has been their manufacturing, distribution, and logistical advantages in domestic and foreign markets (*see* DESIGNING A GLOBAL SUPPLY CHAIN: OPPORTUNITIES AND CHALLENGES; GLOBAL SOURCING STRATEGY: AN EVOLUTION). This has involved (i) creating the appropriate marketing infrastructure "from scratch" (if necessary); as well as (ii) adapting to capitalize on the existing marketing infrastructure in other countries (Craig and Douglas, 2000).

Since international markets vary greatly in terms of existing infrastructure, companies have gone to great lengths to insure consistency in product quality. Through the years, Nestle have invested in systems, equipment, and so on, so that proper production and distribution infrastructure could be put into place that would otherwise not have existed. For example, Nestle devised "milk roads" in China to overcome local deficiencies in transportation and distribution systems.

More often, however, companies have to adapt operations and/or invest in foreign partners in order to succeed abroad. One of the most crucial global investments is the establishment of a reliable distribution system (Arnold, 2000). Companies often differ in their approach to distribution, and the results can be dramatic. For example, Coca-Cola's distribution strategy and ability to adapt to specific regional concerns (e.g., the necessity for vending machines in Japan) has been a key to their global success.

Sometimes, companies mistakenly adapt infrastructure strategies that were critical success factors, only to discover that these changes eroded the brand's competitive advantage. For example, Dell Computer initially abandoned their direct distribution strategy in Europe and instead decided to establish a traditional retailer network through existing channels, with poor results. Ignoring critics who claimed that a direct distribution model would never work in Europe, Dell revamped their direct approach and relaunched their personal computer line with a new management team to execute the direct model that the company had pioneered in the United States, finding greater success as a result.

Developing a proper marketing infrastructure is especially important in developing markets. India still struggles with poor infrastructure and highly restrictive labor laws. Its retail channel structure, although improving, still lags. The quality of public services – education, health, provision of water – is also often lacking. In China, after a series of high-profile product quality scares and crises, government standards were put in place for product quality and safety standards in manufacturing to try to assure overseas consumers and gain their trust.

Right marketing partners. In developing their infrastructure, most global brands have marketing partners of some form in their international markets, ranging from joint venture partners, licensees or franchisees, distributors, ad agencies, and other marketing support personnel. One common reason for establishing brand partnerships is access to distribution. For example, Guinness have very strategically used partnerships to develop markets or provide expertise that the company lacked with their own personnel or capabilities. Partners can also help to make sure supply and distribution chains operate smoothly in different markets. With 226 offices in 70 countries, Seattle-based freight-forwarder Expeditors International help firms keep track of 3000 shipping containers and 2 million pounds of airfreight around the world.

Successful brand partnerships can become key components to overall profitability for each of the parent companies. The value of a partnership can extend far beyond increased sales or access to distribution. Good partners share "brand values" that help maintain brand consistency across markets. For example, McDonald's fierce commitment to product and service standardization is one reason why the retail outlets are so similar all over the world. To achieve such consistency, McDonald's handpick their global partners one by one in order to find "compulsive achievers" who will put forth the desired effort.

Avoiding branding shortcuts. Building a brand in a new international market needs to be done from the "bottom-up" – both strategically and tactically. Strategically, this means concentrating on building awareness first before building the brand image (i.e., to "lay the foundation" for the brand). Tactically, or operationally, this means determining how to best create sources of brand equity in new markets. In other words, the means by which a brand was built in one market (e.g., the particular product, distribution, communication, or pricing strategies and marketing activities) may not be appropriate in another market even if the same overall brand image may be desired.

Many times marketing programs have to be adjusted because the brand is at an earlier stage of development in its new market (*see* MARKET ENTRY AND EXPANSION). In such situations, consumer education about the product itself may need to accompany brand development efforts. When Coca-Cola moved into developing markets in Asia, they encountered consumers who loved the brand, but had never drunk the product before. Not realizing it needed to be chilled, they required education as to the fact that it should be drunk cold.

The recommendation to avoid shortcuts suggests some patience on the part of marketers. Firms may have to "backtrack" to an earlier stage of brand development in these new markets and engage in a set of marketing programs and activities that the brand – in its existing markets – had long since moved beyond. Although the time taken to build the brand in new markets may be compressed because of greater financial resources and a keener understanding of effective strategies and tactics, it could still take some time. The temptation – and often mistake – is to export the current marketing program because it seems to basically "transfer" or "work." Although that may be the case, the fact that a marketing program can meet with acceptance or even some success in a market does not mean that it is the proper marketing activity in terms of building a strong, sustainable brand equity there.

For example, when Nike made a big push into Europe in the early 1990s, they were too aggressive in their approach and overrelied on their well-known American sports stars. Although athletes such as Michael Jordan, Bo Jackson, and Wayne Gretzky were known to varying degrees in Europe, they represented sports (i.e., basketball, football, baseball, and hockey) that were not as popular in Europe as they were in America. As a result, the ads that were so captivating in

the United States generated much less fanfare in Europe. The brand met with some success, but failed to live up to its potential. Nike management soon came to realize that Nike's brand mantra of "authentic athletic performance" had a different meaning in the European culture. Instead of using American heroes playing American sports, Nike adopted a more "grassroots approach" to better reflect authentic athletic performance "European style." Soccer, or football, thus became an indispensable ingredient, and Nike sales began to rise accordingly.

Not taking shortcuts helps to create marketing momentum and support from a growing customer base in the new market. Red Bull deliberately adopted a phased roll-out program in entering a new market with imposed scarcity to help drive up interest and demand in their new functional energy drink product. Jamaica-based Digicel have successfully conquered many politically unstable third-world countries such as Papua New Guinea, Haiti, and Tonga by developing product and programs that appeal to the poor consumers who are typically otherwise overlooked. The fierce loyalty of these consumers helps to protect Digicel from any overly aggressive government actions or interventions.

Strike a Balance in Global Brand Management

Ideally, a single marketing program could be devised and implemented that would turn out to be the most effective and efficient possible option for each and every country in which the brand is to be sold. There are many advantages to launching such a globally consistent marketing program for a brand: economies of scale in production and distribution; lower marketing costs; power and scope; consistency in brand image; ability to leverage good ideas quickly and efficiently; and uniformity of marketing practices.

Unfortunately, such uniformly optimal global marketing programs are rarely possible. One implication of the similarities and differences across international markets is the need to blend local and global elements in marketing programs. The best examples of global brands retain a thematic consistency and alter specific elements of the marketing mix in accordance with consumer behavior and the competitive situation in each country. An oft-heard – and sometimes modified – expression of prescriptive advice to marketers of global brands is to "Think Global, Act Local." In that spirit, HSBC are even explicitly positioned as "The World's Local Bank."

Perhaps the quintessential global brand, Coca-Cola, deliberately keep the basic look and packaging of their Coke brand the same everywhere (except in countries where laws dictate use of local language). Yet, the company simultaneously stress that the brand be *relevant* and well positioned *relative* to competition in different markets. They use different advertising agencies in different countries in order to make the brand feel local and be well positioned relative to local competition. For example, in Australia the advertising appeals to the same "classic, original" ideals but in a very Australian fashion. As a result, Coke becomes entwined with the cultural fabric of the country, just as it has in the United States. Over time, this yields an advantage with younger generations who do not even think of Coke as an imported brand. An illustrative example that Coca Cola recount is of a Japanese family visiting the United States for the first time: the young son, upon passing a vending machine, joyfully exclaimed to his parents, "Look, they have Coke here too!"

Most brands are adapted to some extent to reflect differences in consumer behavior, brand development, competitive forces, and the legal or political environment across markets (*see* STANDARDIZATION/ADAPTATION OF INTERNATIONAL MARKETING STRATEGY). Even global brands undergo some changes in product features, packaging, channels, pricing, or communications in different global markets. Some aspects of the marketing program tend to be adjusted less frequently or dramatically than others (e.g., brand elements such as brand names, logos, packaging, and signage) whereas others are adjusted more frequently or vary more dramatically across markets (e.g., advertising and other communications and especially distribution channels). Even something as simple as a brand name may involve various decisions in terms of the use of dual names, choices between phonetic or

semantic translations, and so on (Zhang and Schmitt, 2001; Hong *et al.*, 2002).

Even in advertising, it is not uncommon to have the same creative theme globally, but to adapt the specific execution to appropriate local markets (*see* INTERNATIONAL ADVERTISING – IS THERE STILL A STANDARDIZATION VERSUS LOCAL ADAPTATION DEBATE?). Apple Computer's "Mac vs. PC", which was voted the best US ad campaign of 2006 by *Adweek* magazine, features two actors bantering. One is hip-looking (Apple), and the other is nerdy-looking (PC). Apple dubbed the ads for Spain, France, Germany, and Italy, but chose to reshoot and rescript for the United Kingdom and Japan – two important markets with unique advertising and comedy cultures. The United Kingdom followed a similar formula but tweaked the jokes to reflect British humor; the Japanese ads avoided direct comparisons and were more subtle in tone. GE's "Ecomagination" ad campaign retains the same message globally, but substitutes different ad creative and executions in Asia and the Middle East to reflect the cultural interest there.

The challenge, of course, is to determine the nature of the balance in the marketing program – which elements to customize or adapt and which to standardize. This balance can depend on a host of factors, for example, the cultural flavor of the brand or unique characteristics of the market in question. Customization may imply adjusting some aspect of the marketing program and/or the desired brand image (e.g., by the creation or deletion of brand associations). In some cases, because of differences in consumer behavior or because of historical market factors, brand positions may fundamentally be different in different markets (Aaker and Joachimsthaler, 1999).

- Heineken beer is a high-end superpremium offering in the United States, but more middle-of-the-road in their Dutch home market.
- The domestic image of Honda automobiles in Japan, on the other hand, is richer and the brand is more strongly associated with speed, youth, and energy than in overseas markets such as the United States, where it is seen as a reliable, quality vehicle.
- The Toyota Camry is the quintessential middle class car in the United States, but is positioned at the high end in China, even though the cars in the two markets differ only in very cosmetic ways.

Even if the positioning does not vary for a brand, the specific products and services that are emphasized may still vary. IBM have a two track approach with their services business: in the United States, where clients often are economizing, they focus on helping with cost cutting; whereas in developing markets, where clients are seeking to modernize and catch up with other countries, they help customers with their technology infrastructure.

Not sufficiently adjusting marketing for a global brand in a new market can lead to negative consequences. Much of the success of Finland's Nokia in cell phones has been due to their intense focus on innovation, design, and engineering in producing a wide range of products that vary in quality, price, and so on. This range has allowed them to be successful in both developing and developed markets. Despite this global success, however, they found their market share slipping in the United States when they failed to customize their cell phones to the various wireless carriers. Improving their partnership with wireless carriers resulted in greater retail presence in showrooms and stores, a critical success factor in that market.

As another example, despite some financial success, Korea's LG decided to hire a number of top executives from Western firms to help transform themselves from what they saw as "an engineering powerhouse that excelled in manufacturing and selling in different parts of the world" to a "globally efficient, trend-setting organization." The new executives were charged with standardizing the hodgepodge of processes and systems that LG had developed in different markets in purchasing, the supply chain, marketing, and other areas. In particular, a single agency (London's Bartle, Bogle, and Hegarty) was given global responsibility in marketing to sell an increasing number of higher-end products.

Finally, another issue in adaptation is the country of origin (*see* "COUNTRY OF ORIGIN" AS BRAND ELEMENT). Country of origin can

clearly play a different role in positioning brands in foreign versus domestic markets. Domestically, country-of-origin associations may attempt to stir feelings of patriotism and cultural heritage, but in foreign markets, it may be a means to leverage existing perceptions and beliefs about the home country. Even a perception of globalness for a brand – by sending a quality signal, tapping into cultural myths, and reinforcing a sense of social responsibility – can improve brand evaluations (Holt *et al.*, 2004; Steenkamp *et al.*, 2003).

There are numerous examples of companies that leverage country-of-origin associations to help position their brands globally: "German engineering" for BMW and Mercedes, "Aussie good cheer" for Foster's beer, and "French elegance" for Chanel, Dior, and Louis Vuitton. A Western image can be helpful in developing markets, as Coca-Cola discovered in China (Batra *et al.*, 2000). Part of their success against local cola brand Jianlibao was due to the brand's symbolic values of modernity and affluence. But not all country images are positive, and country-of-origin effects can be tricky when complex multicountry component and assembly are involved, posing challenges to global marketers (Tse and Lee, 1993).

Balancing local and global control. A key theme with organizational structures, entry strategies, and coordination processes and mechanisms for global brands is the need to balance global and local control. Coca-Cola, for example, have distinguished between marketing activities that would appear to dilute brand equity from those that would not appear to be as efficacious as desired. Headquarters would stop the former from occurring, but would not stop the latter. They would leave it to the local manager's judgment as to the activity's appropriateness, but also hold him or her responsible for its success or failure.

Similarly, Levi's have balanced global and local control with a "thermometer" model. Marketing elements below the "freezing point" are fixed: "brand soul" or essence and logos are standardized worldwide. Above the freezing point, product quality, pricing, advertising, distribution, and promotions are all "fluid," meaning each international division can handle the marketing mix elements in any way that they feel is appropriate for their region.

McDonald's also have allowed countries and regions more latitude to customize around their basic layout and menu staples. In China, they substitute corn for fries in Happy Meals, some stores in United States blend fruit smoothies, and Australia and France have Starbucks-like lounges. In cities plagued with horrendous traffic problems – for example, Manila, Taipei, Jakarta, and Cairo – they actually deliver McDonald's meals via fleets of motor scooters.

Brand architecture. Companies such as Kraft, Unilever, and P&G market global brands in countries all over the world, but they also have a number of local brands or "local treasures" as they sometimes call them, which have built up strongly resonant consumer franchises (Schuiling and Kapferer, 2004). Unilever's local jewels, for example, include Ben & Jerry's ice cream, Suave hair care, and Wish-Bone salad dressings in the United States; and Bovril and Marmite spreads, Peperami spicy salami, and Pot Noodle noodle snack foods in the United Kingdom. For companies with such varied brand portfolios (*see* MANAGING THE GLOBAL PRODUCT PORTFOLIO), their challenge in many cases is to ensure that the global brands stay relevant in local markets and, on the other hand, local brands are able to compete on a global stage in their home market or potentially even elsewhere (Douglas *et al.*, 2001).

Nestle are a company that have benefited significantly from a balance of global and local control. Some decisions, such as branding, follow strict corporate guidelines. The company have 10 *worldwide corporate* strategic brands, including Nestle, Nescafe, Maggi, and Carnation. There are 45 different strategic *worldwide product* brands, including Kit Kat, Coffeemate, and Crunch. There are 25 *regional corporate* strategic brands, including Perugina, Findus, and Stouffer's. There are 100 *regional product* brands, including Eskimo, Taster's Choice, and Go-Cat. Finally, there are 700 *local strategic* brands that are important to particular countries, including Brigadeiro in Brazil.

Most other decisions, however, are primarily decided upon by the local managers. Nestle's policy is to recognize that they are foreigners in

any country outside Switzerland and to simply let the local managers run most of the business operations. For example, the company do not do corporate-level strategic planning. Instead, the top managers in each region tell headquarters what they plan to do and a combination of bottom-up or top-down approaches is used to finalize a strategy. Headquarters meet once a year with each of the country managers to discuss strategic issues. The bottom line resides with headquarters, and they have the power to force the local managers to adopt a policy, if necessary. This happened with ice cream in the United States.

The country manager classified the product as a dairy product, whereas headquarters viewed it more strategically as a frozen confectionery. The country manager proposed machines and cones, and headquarters countered with self-manufacture and direct store delivery. Today, Nestle are a strong number two in the ice cream impulse segment, and have invested in a 17% stake in Dreyer.

Establish operable guidelines. Brand definition and guidelines must be established, communicated, and properly enforced so that marketers in different regions have a good understanding of what they are expected to do and not to do. The goal is to clearly set the rules for how the brand should be positioned and marketed. Hence, everyone within the organization understands the brand's meaning and can translate that meaning to satisfy local consumer preferences.

Aaker and Joachimsthaler (1999) put it this way: "Global brand leadership means using organizational structures, processes, and cultures to allocate brand-building resources globally, to create global synergies, and to develop a global brand strategy that coordinates and leverages country brand strategies." Specifically, they advocate four key ideas:

- stimulate the sharing of insights and best practices across countries;
- support a common global brand-planning process;
- assign managerial responsibility for brands in order to create cross-country synergies and to fight local bias; and
- execute brilliant brand-building strategies.

Brand definition and communication is aided with some sort of document that details what the brand is and what the brand is not. In terms of brand documentation, Coca-Cola have a strategy document that clearly articulates their strategy and how the brand positioning is manifested in various aspects of the marketing mix elements. This document sets out the parameters for the brand and, therefore, determines how much is left to chance. Similarly, McDonald's operating manual imposes rigorous worldwide controls (e.g., the 19 steps to cook and bag French fries). Nestle ensure that branding decisions at least follow strict corporate guidelines.

Brand mantras may be especially helpful in providing easily understood brand guidelines. A *brand mantra* is an articulation of the "heart and soul" of the brand. Brand mantras are short, three- to five-word phrases that capture the irrefutable essence or spirit of the brand positioning. Their purpose is to ensure that all employees within the organization and all external marketing partners understand how the brand should be projected to the consumers so that employees and partners can adjust their actions accordingly. As noted above, Nike's brand mantra is "authentic, athletic performance" and effectively translating that in different markets has been one of their keys to global success.

Disney's brand mantra has been "fun, family entertainment." To establish global guidelines for licensing, Disney assign them to one of the three categories: (i) acceptable to license without permission (e.g., T-shirts); (ii) not permissible to ever license (e.g., toilet paper); and (iii) requires validation from headquarters to license (about 20 categories – e.g., air fresheners). Internationally, Disney noticed that these "gray areas" grew bigger and more numerous. The company also try to identify those product groups that may be more amenable to localizing than others. For example, movies cannot be tailored for the European market because it is difficult to determine what will be attractive to those consumers. On the other hand, certain items in the Disney store may sell well in Germany but not in Japan.

Finally, for all of this to work, there must be effective lines of communication and means to transfer knowledge within and across regions. Coca-Cola stress the importance of

having people "on the ground" who can effectively manage the brand in concert with their headquarters in Atlanta. For example, to facilitate coordination, much training occurs in headquarters, a sophisticated e-mail and voice mail system is in place, and global databases are available. The goal of this heavily integrated information system is to facilitate the local manager's ability to tap into what constitutes "relevance" in any particular country and then communicate those ideals to headquarters or other parts of the region. Coca-Cola are thus able to transfer product ideas within and across regions. Japan, long a hot-bed for new product ideas, has been joined by other Asian markets such as China as the company expands their product portfolio to include water, juice, tea, coffee, and even "gel" beverages. From China came Minute Maid Pulpy; from Hong Kong came Nestea with Aloe Pulp.

Effectively transferring successful marketing ideas from one region to another region is a key priority for many firms. Rather than developing global products for jointly owned Renault and Nissan, CEO Carlos Ghosn has mandated that companies design for local tastes and have the flexibility to export the design to other regions to tap into similar consumer trends. As an example, the no-frills Logan was developed by Renault for Eastern Europe and Latin America but it found another home in France. Along with the specific products that cross a region, ideas and a way of thinking may also transfer in the process. Ghosn teamed Nissan and Renault with Bajaj Auto to sell a $3000 car in the Indian market in part to infuse those companies with India's low-cost design thinking: "They understand frugal engineering, which is something we aren't as good at in Europe or Japan."

Conclusions

Increasingly, it is imperative that marketers properly define and implement a global branding strategy. Although global branding offers many potential benefits to a firm, there are also significant risks. Despite these potential pitfalls, a number of marketing pioneers have successfully established global brand powerhouses over the last decade or so. On the basis of the experiences and practices of leading global brands and the learning from academic research, this article has identified three key steps to building a successful global brand: (i) understand the global consumer context; (ii) build proper global marketing infrastructure; and (iii) strike a balance in global brand management. Although building a strong global brand can involve a number of complex issues and considerations, successfully accomplishing these three steps can increase the odds of success.

Bibliography

Aaker, D.A. and Joachimsthaler, E. (1999) The lure of global branding. *Harvard Business Review*, **37**, 137–144.

Arnold, D. (2000) Seven rules of international distribution. *Harvard Business Review*, **38**, 131–137.

Batra, R., Ramaswamy, V., Alden, D.L. *et al.* (2000) Effects of brand local and nonlocal origin on consumer attitudes in developing countries. *Journal of Consumer Psychology*, **9**, 83–95.

Craig, C.S. and Douglas, S.P. (2000) Configural advantage in global markets. *Journal of International Marketing*, **8**, 6–26.

Douglas, S.P., Craig, C.S., and Nijssen, E.J. (2001) Integrating branding strategy across markets: building international brand architecture. *Journal of International Marketing*, **9**, 97–114.

Hofstede, G. (1980) *Culture's Consequences*, Sage Publications, Beverley Hills.

Hollis, N. (2008) *The Global Brand*, Palgrave Macmillan, New York.

Holt, D.B., Quelch, J.A., and Taylor, E.L. (2004) How global brands compete. *Harvard Business Review*, **82**(10), 68–75.

Hong, F.C., Pecotich, A., and Schultz, C.J. II (2002) Language constraints, product attributes, and consumer perceptions in east and southeast Asia. *Journal of International Marketing*, **10**, 29–45.

Johansson, J.K. (2009) *Global Marketing*, 5th edn, McGraw-Hill, New York.

Khanna, T. and Palepu, K.G. (2006) Emerging giants: building world-class companies in developing countries. *Harvard Business Review*, **84**(10), 60–69.

Mahajan, V., De Moraes, M.V.P., and Wind, J. (2000) The invisible global market. *Marketing Management*, **9**(4), 31–35.

Schuiling, I. and Kapferer, J.-N. (2004) Real differences between local and international brands: strategic implications for international marketers. *Journal of International Marketing*, **12**, 97–112.

Steenkamp, J.-B.E.M., Batra, R., and Alden, D.L. (2003) How perceived brand globalness creates brand value. *Journal of International Business Studies*, **34**, 53–65.

Tse, D.K. and Lee, W. (1993) Removing negative country images: effects of decomposition, branding, and product experience. *Journal of International Marketing*, **1**, 25–48.

Zhang, S. and Schmitt, B.H. (2001) Creating local brands in multilingual markets. *Journal of Marketing Research*, **38**, 313–325.

base of the pyramid markets: culture insights and marketing implications

Cheryl Nakata

Introduction

Much of the world's population is poor. According to the World Bank, 2.7 billion persons, representing over half of the world's population, live on less than $2 a day. Within that group, 1.1 billion subsist on just $1 per day. The greatest concentrations of poverty occur in developing countries and regions, such as China (596 million) and sub-Saharan Africa (514 million) (*see* EMERGING MARKETS). As a market, the poor in these geographies – termed the "base of the pyramid" or BOP – has been largely ignored (Prahalad, 2005).[1] Businesses have focused instead on the middle and top of the pyramid, selling products to wealthier consumers in developed or developing countries.

However, the immense size of the BOP market and relative dearth of competition for its loyalty spell significant growth opportunities for global companies (*see* GLOBAL MARKETING STRATEGY). Firms are beginning to seize these opportunities, specifically by creating new, or significantly adapting existing, goods and services for the unique needs of BOP consumers. Examples are Hindustan Lever, which developed a disinfectant soap for villagers in India to reduce widespread diarrheal disease; Haier, which designed a small washing machine priced at $37 for first-time white goods buyers in China; and Nokia, which produced a $25 cell phone, featuring a speaking alarm clock and an iconic address book, for illiterate Brazilians. These firms are hoping to not only generate significant new revenues, but also fulfill a mandate for greater corporate social responsibility. By providing the poor with access to quality, value-laden products and services, such enterprises aim to open the door of full participation to persons historically excluded from the global marketplace (*see* GLOBAL MARKETING ETHICS).

Despite the opportunities, most businesses find the idea of tackling the BOP market, with all its attendant challenges, daunting. Buyers in this market are characterized as largely illiterate, in poor health, of meager resources, inaccessible by media, geographically isolated, and ignorant about consumption (Prahalad, 2005). Marketing to such buyers requires deep expertise to work within these constraints. A particular need is cultural understanding (*see* SOCIETY, CULTURE, AND GLOBAL CONSUMER CULTURE). Compared to the top or middle of the pyramid, the BOP is culturally remote to businesses and their managers, who are generally unfamiliar with and have not experienced the severe and constant deprivations of subsistence living. These deprivations result in a distinctive way of life, that is, a culture that affects consumption attitudes and practices (Viswanathan, 2007). Effectively addressing the BOP market, then, rests on understanding the BOP culture.

The purpose of this article is to provide insights on the BOP culture and explore implications for firms marketing to the BOP, perhaps the last, largest, and most difficult market frontier. Because business interest in the BOP is emergent, it may be opportune to elaborate on cultural matters, informing and guiding future strategies and approaches.

A Culture of Poverty

From the perspective of outsiders, BOP consumers live a dark and hopeless existence. Grinding poverty means dehumanizing routines for survival, such as scavenging trash hills for tossed food and foregoing medications when desperately ill (Hill, 2002; Viswanathan, 2007). Threats to shelter, security, and all forms of well-being are unrelentingly harsh. Fifty years ago, Lewis (1959) provided an indelible portrait of life at the bottom titled *Five Families: Mexican Case Studies in the Culture of Poverty*. The ethnography vividly described the daily struggles of the poor. More importantly, Lewis articulated and received wide attention for his treatise that the poor possess a unique culture. According to Lewis, the poor respond to material scarcities with negative attitudes and behaviors, such as alienation and crime. These habits of thought and action are rooted in early childhood and transferred within families (Lewis, 1966, p. xiv):

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

> The culture of poverty is not only an adaptation to a set of objective conditions of the larger society. Once it comes into existence it tends to perpetuate itself from generation to generation because of its effect on the children. By the time slum children are age six or seven they have usually absorbed the basic values and attitudes of their subculture and are not psychologically geared to take full advantage of changing conditions or increased opportunities which may occur in their lifetime.

Lewis' view has been highly influential in formulating public and welfare policies, as well as popularizing the notion of a unique culture of poverty. The question raised is, whether the BOP has a "culture of poverty" as described by Lewis? The argument here is that while the BOP may have distinctive cultural characteristics, these characteristics are not entirely negative or as deterministic as portrayed by Lewis. Critics of Lewis point out that a subtle but important precept of Lewis' theory is that the poor exist within a self-defeating culture of their own making (Leacock, 1971, pp. 7–37): they are unwilling and/or unable to pull themselves up by their bootstraps to plan for, work toward, and live a better future. Low aspirations, weak self-discipline, and high social exclusion combine to doom them to short, hellish existences.

Studies suggest – and the counterview is forwarded here – that while the poor face extraordinarily difficult circumstances, they desire and dream no less than the nonpoor for better lives. BOP persons want more like everyone else. Critically, if given the chance to save money, start businesses, educate their children, care for their physical health, and lead productive lives in civil harmony with others, they will do so (Hill, 2002; Jones and Lou, 1999; Viswanathan, 2007). The difference is that externalities, experiences, knowledge, and resources limit what they do achieve. In other words, the poor do not inherently lack moral fortitude or agency to change their lives; instead, what they lack are the means and opportunities to live differently. The choices available to them are fewer and thus their decisions differ, all of which result in a culture that is more complex and dynamic as well as less predictable and pathological than Lewis suggested.

A Culture with Paradox

The temptation in broaching the BOP, as with any culture, is to assign it a monolithic quality. "All of the poor are _____" (fill in the blank) captures the sentiment. However, the BOP cannot be so simplistically characterized. Within the culture of poverty – consistent with its complex and dynamic nature – there is also a major paradox. The argument is that the BOP is simultaneously *both* poor and rich, and specifically that BOP consumers are deprived in very fundamental ways yet endowed with other resources. It is argued that the deprivations of the poor, which are deep and dehumanizing, should not be dismissed. This assertion of paradox has been made to encourage a deeper look at the market.

The BOP has little economically. This condition is obvious. What is less apparent is that the deprivation is accompanied by many others. As the Nobel Laureate Sen (1999) observed, poverty is not just the scarcity of economic means; it is more broadly and subtly multiple forms of deprivation or "unfreedoms." Among these unfreedoms are social isolation due to the stigma of poverty, physical illness from unsanitary living conditions, psychological distress through constant threats of violence, low employment because of limited schooling, and fatigue from labor-intensive drudgery (e.g., fetching water for hours each day because there is no indoor plumbing or a nearby drinkable water source). Even exclusion from the marketplace, or not having access to quality goods within one's means, is a deprivation. Additionally, the poor live under conditions – economic, sociocultural, and political-regulatory – that disadvantage them relative to their nonpoor counterparts. Minimal job opportunities, inadequate housing, atomized retailing, media darkness, high prices, and low choice, for instance, are typical of impoverished communities.

Importantly, too, BOP individuals suffer from associated emotional and cognitive stresses. Because of severely restricted incomes, they hold onto possessions that would be considered unacceptable by others who are better off (e.g., having only one pair of worn and tattered jeans), and do not buy items most others own and take for granted (e.g., new pencils at the start of the

school year). The comparative lacks generate emotions of rage, fear, and hopelessness (Hill, 2002). Low literacy, numeracy, and education also contribute to cognitive difficulties in making informed purchasing choices. For example, the inability to calculate price per ounce or read special promotion signs leads to buying heuristics such as always getting the same product, looking for the smallest size regardless of price, and focusing on a single product attribute rather than trading off among multiple attributes (Viswanathan, Rosa, and Harris, 2005). The poor are aware and ashamed of their limited "capabilities" in Sen's parlance, creating stressful shopping and consumption experiences.

At the same time, such lacks stand in stark contrast to the considerable internal and external resources of the BOP. BOP persons have been observed to be enormously creative given the constraints they live under (Subrahmanyan and Gomez-Arias, 2008). Their inventiveness reflects an internal fortitude or resilience in the face of significant challenges (Sen, 1999). For instance, while their purchases may not be optimal in the environments of abundance surrounding middle- or upper-income consumers, BOP individuals have developed sophisticated buy and use strategies that imaginatively stretch their budgets in their context of scarcity. Typically, BOP consumers live in fragmented retailing environments, so there are only a few stores nearby carrying small inventories of limited product choice. Even in these conditions, BOP consumers manage to negotiate purchases with retailers, demanding and receiving quality products contrary to the belief that the poor buy only the cheapest goods (Viswanathan, 2007). Along with such internal resources, these consumers have external resources that enable them to know of and obtain what they need. Their communities are characterized by tight social networks leveraged for this purpose (Subrahmanyan and Gomez-Arias, 2008; Viswanathan, 2007). Families, neighbors, and friends help them make purchases they may not be able to afford on their own. And small retailers, with whom they have forged strong ties in rural settings, provide credit and customize purchases where no alternatives (e.g., banks or wide-assortment stores) exist. In short, the BOP culture is characterized by a paradox of enormous scarcity on the one hand and an abundance of social, noneconomic capital on the other.

A Culture Embedded and Shaped

Another aspect of the BOP culture is that it is embedded in and shaped by other cultural forces, specifically, consumer culture, national culture, and globalization, as depicted in Figure 1. Consumer culture refers to shared meanings and unifying values tied to consumption (*see* SOCIETY, CULTURE, AND GLOBAL CONSUMER CULTURE; GLOBAL CONSUMERISM AND CONSUMPTION). While the consumer culture is more nascent in developing countries relative to industrialized settings, it, nonetheless, presents and projects values, aspirations, and behaviors whereby individuals formulate and act upon social identities through consumption. That the consumer culture exists in subsistence economies like Sri Lanka's and Turkey's is evident in the awareness and purchase of global brands such as Coca-Cola and Nestle even among persons who supposedly cannot afford them (Ger, 1997). The goal of marketers is to develop a strong consumer culture through which the BOP culture is encouraged to search for and purchase particular products and services. In other words, the consumer culture helps shape BOP culture. Over time, the BOP culture likewise influences consumer culture.

An example of this dynamic is Nokia's marketing efforts in India, which has altered the local BOP culture. India is the fastest growing and third-largest mobile-phone market in the world with approximately 220 million users. Nokia aggressively entered the market several years ago and captured the lion's share of 60% by introducing basic phones priced affordably for the low-income consumers (around $20) as well by offering full-featured, more expensive models for higher-income individuals (upwards of $1200). Accompanied by broad distribution and heavy brand advertising, Nokia phones have become pervasive throughout India, including in poor rural villages. Besides low pricing, Nokia further addressed the needs of rural villagers by adding a flashlight or torch feature to its phones,

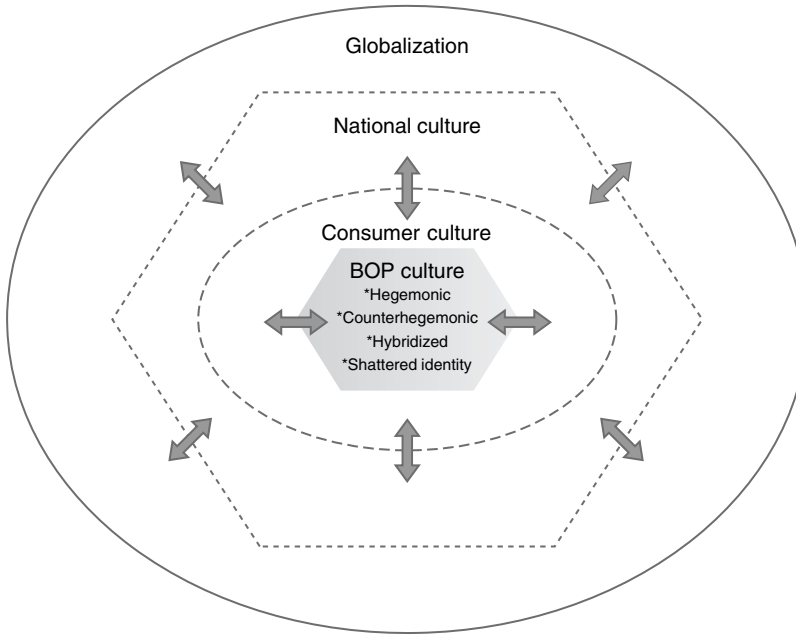


Figure 1 BOP culture as embedded, shaped, and heterogeneous.

given that electrical lighting is spotty and unreliable in these areas. Nokia's efforts to create a consumer culture have paid off handsomely in terms of high demand for its mobile phones, even among the poor. As a consequence, the BOP culture in India has been transformed. The poor, or BOP, consumers have embraced the phones and incorporated them into their daily lives, fundamentally altering how they socialize, do business, entertain themselves, and communicate with one another (www.ibef.org).

The second force affecting the BOP culture is national culture. National culture reflects the customs and traditions of a country. While no national culture is completely consistent or integrated, there are some shared patterns of heart and mind that influence a country's citizens (Hofstede, 2001). Through religious, political, and educational institutions, national culture exerts an influence on and is, in turn, influenced by consumer culture. This means that the BOP culture is indirectly affected by national culture, reflecting a complex layering of cultural interactions. An example of such interactions is given by Viswanathan (2007), who describes how the poor in some societies avoid taking on debt to make purchases because of the norm of face loss or social embarrassment associated with credit. Thus national culture, in the form of face saving, affects the consumer and BOP cultures. It affects both cultures by lowering the desirability of making purchases without cash in hand, and encouraging other means to engage in purchase behaviors, including borrowing funds from friends, families, and neighbors. Furthermore, Viswanathan points out how collectivism, a national culture value, injects itself in the marketing and consumption practices of these poor communities. Collectivism fuels word-of-mouth effects among consumers, strengthens ties between marketers (especially retailers) and customers, and facilitates group-based financing, purchasing, and use of products that would not be affordable on an individual basis. An illustration of the latter is the pooling of monies on a monthly basis among community members to provide a loan to one participant for a major purchase or to support a business, with the understanding that the monies are to be repaid and then used to provide a loan to another member the following month.

Finally, the third cultural force affecting the BOP culture is globalization. Globalization creates shifts – sometimes quickly, sometimes slowly – in the social, economic, political, and technological landscapes of developing countries. Globalization therefore also affects BOP populations and their culture within these countries. For instance, globalization (specifically, its promise of industrialization and economic growth) can motivate a government to liberalize its trade policies by permitting the wider importation of foreign goods and advertisements, as has occurred in China, Brazil, and Indonesia, among others. BOP consumers are subsequently exposed to new goods and promotion messages, being informed about the availability of, and offered perhaps for the first time, branded detergents, ice cream, washing machines, and other products of good quality that are taken for granted by the more affluent. Through continual exposure to such items and advertisements, these consumers individually and their BOP culture collectively are altered, both in terms of product purchases and usages as well as beliefs about the nature, quality, and goals of their lifestyles as reflected in consumption choices. As noted by Ger, with respect to developing countries (Ger, 1997, pp. 110–111), "Because development includes modernization and marketization, the increased availability, display, and advertising of mostly foreign products fuels aspirations for the good life and raise consumption expectations . . . the image of the good life in less affluent countries is one of being a successful participant in a consumption-oriented society . . . There is widespread desire for the goods that prominently surround people in the less affluent world."

A Culture Heterogeneous

As a consequence of embeddedness in consumer culture, national culture, and globalization, the BOP culture is actively under negotiation. Not all the forces encompassing and shaping the BOP culture move in the same direction, and the BOP culture has its own pulls and preferences. Thus it represents an area of contested change. For instance, globalization may promote the acceptance of foreign goods by the BOP, but national culture may do the opposite, fanning the flames of nationalism and rejecting the encroachment of alien influences. The BOP culture is caught in the middle, subject to conflicting tensions. The term *negotiation* means that the BOP culture is in a fluid state, with a heterogeneity of coexisting meanings and SUBCULTURES. These heterogeneous meanings and subcultures reflect a range of reactions to and engagements with the consumer culture, national culture, and globalization. More specifically, following Askegaard, Arnould, and Kjeldgaard (2005) and Ustuner and Holt (2007), it is proposed that those meanings/subcultures fall into four categories: hegemonic, counterhegemonic, hybridized, and shattered identity. These categories are also represented in Figure 1.

When consumers encounter a culture not their own, such as by migrating from a traditional rural setting into a contemporary urban city – as has occurred to 1 billion transplants throughout the developing world, representing the largest human migration in modern history – they develop identity projects that help them navigate the contradictory discourses (*see* GLOBAL CONSUMERISM AND CONSUMPTION). For some, navigation is an attempt to assimilate the other culture, embracing it while rejecting their own. This is the termed the *hegemonic* response, whereby the other, now surrounding or dominant, culture is taken and absorbed. It is truly a project of assimilation in that it entails discarding habits of thought, feeling, and action, and concomitantly acquiring unfamiliar ones. An example of a hegemonic response is Penaloza's (1994) study of Mexican American immigrants, some of whom took on American ways of life while abandoning their more traditional and less consumption-oriented culture.

A second response is one of resistance. Termed *counterhegemonic*, individuals attempt to retain their culture of origin, reinterpreting it yet staying true to its moorings, within the new culture. This response is primarily one of putting at bay the other culture, though it is not pure rejection in that a few selected symbols of the other culture are adopted. An example given by Ustuner and Holt (2007) is of Turkish village women who move into squatter enclaves in a major Turkish city. They purchase TVs, video

cameras, and other Western goods while holding onto a "country bumpkin" lifestyle, such as by continuing to eat meals on the floor as they did back "home." Importantly, even these symbols of the other culture are used in ways consistent with, and reshaped into, the old or original culture. Thus, for example, freezers are shared. A whole cow is slaughtered, but then each family's share of meat is stored in the freezer, all in keeping with traditional communal practices. In essence, though deterritorialized (i.e., having moved physically away from their villages), the women reestablish their rural habitus in the urban context.

A third response is termed *hybridization*. Here, two or more cultures are truly merged to create a third, new culture. Consumers mix and match identity positions, interpreting and making use of majority and minority discourses. This is an existentially challenging endeavor, but can be managed by some persons with little difficulty. A BOP consumer taking this route could, for example, bring together elements of his/her native local culture, for example, habits of food preparation from the rural community, with those of the national culture, for example, ways of dressing from the capital metropolis.

Finally, when a new culture is aspired to while the old is rejected, but the individual experiences neither as fully his or her own, perhaps as a consequence of class distinctions too great to traverse (possessing inadequate social, cultural, and economic capital to make the transition), the result is termed a *shattered identity* project. This state is experienced as one of futility, as the individual lives unhappily in the nether world between two cultures. Ustuner and Holt (2007) provide rich descriptions of such a state among young women who come from villages and aspire to the urban life of ease and abundance. However, failed attempts to marry outside their relatively poor and uneducated circles results in having to remain in their squatter communities, dejected about their current lives and futures.

Together, these four categories represent the spectrum of possible responses by the BOP in relation to its cultural milieu. The categories have been put forth to point out that the BOP culture is not uniform, and that any acculturation enterprise it undergoes is a rich, nuanced process with varied forms of adaptation.

Marketing Implications

In view of the above, there are several implications for marketers. One implication is that since the BOP does not differ in life aspirations so much as opportunities with the middle and top of the pyramid, per the revised poverty of culture thesis offered here, marketing strategies and tactics that appeal to these shared aspirations, such as the desire for a happy, healthy, and good life, should be effective with BOP consumers. At the same time, the restricted opportunities faced by these buyers must be addressed so that the goods and services are truly accessible, not offered as tantalizing yet unattainable prospects (*see* ETHICAL MARKETING AND MARKETING STRATEGY). The latter only serves to frustrate and demoralize the poor. Forms of accessibility include lowering prices so the items are affordable, and distributing products within easy, nearby reach in rural areas. Prahalad (2005) provides case studies of firms that have done both using highly inventive means.

A second implication, which stems from the insight that the BOP is a culture of paradox, is to apply marketing approaches that leverage the strengths or resources of this population while circumventing the weaknesses or lacks. As noted earlier, BOP communities tend to be rich in social capital. Through strong ties with families, friends, neighbors, shopkeepers, and others within their communities, BOP consumers obtain information and other aid in making purchases. These social resources can be leveraged in marketing, such as by relying on word-of-mouth (rather than mass) advertising; incenting retailers in personal (rather than impersonal) sales; providing collective, for example, family-based (rather than individual) forms of credit and purchasing; and organizing direct (rather than indirect) sales and distribution systems where relation marketing methods are effective. In terms of weaknesses or lacks, it has been noted that BOP individuals are hampered by a host of deprivations, among them low literacy. To help ameliorate the literacy issue, a marketing solution is to design packaging and advertising that is pictorial for ease of comprehension and, at the same time, emphasizes concrete attributes of the product or service (Viswanathan, Rosa,

and Harris, 2005). Other limitations in BOP capabilities also need to be accounted for in marketing programs to ensure their efficacy.

With respect to the embedded nature of the BOP culture, an implication is to conduct deep, not superficial, market research to fully understand the culture's multiplex and evolving nature (*see* ETHNOGRAPIC RESEARCH). The encircling forces of consumer culture, national culture, and globalization impact the BOP culture, resulting in a unique, changing sociology. Figure 1 underscores that the BOP culture is best understood when contextualized. Hence each BOP market deserves firsthand examination, and should be periodically restudied to develop appropriate, contemporary marketing strategies. As an illustration, businesses engaged in BOP markets such as Procter and Gamble are finding that immersion of their staff in these communities, that is, having staff live for extended periods with poor families, is critical for obtaining insights on how best to meet their needs. While learning can be carried over from one BOP market to the next, the assumption that they are all the same and should be treated as such is as likely to be erroneous as believing all top of the pyramid markets are identical. Importantly, as well, since learning about BOP markets in general is at its infancy, now is the time to invest in significant primary market research.

A final marketing implication, which results from the insight that the BOP culture is under negotiation, is that segmentation and targeting are useful (*see* MARKET SEGMENTATION AND TARGETING). There may be at least four subcultures within the BOP, each of which has a distinct view and set of behaviors in relation to goods and services representing other cultures: hegemonic, counterhegemonic, hybridized, and shattered identity. The first three are to varying degrees adoptive of such goods and services, with their reasons and ways of doing so differing. Astute marketers will recognize these distinctions, segmenting the market instead of treating it as a homogeneous whole, selecting the segments more amenable to their overtures, and then customizing the marketing mix accordingly. Marketers may discover en route that there are other segments, and pursue ones that are more receptive to their companies' particular offerings. In conclusion, the BOP market poses enormous opportunities as well as challenges for global businesses. A chief need in tackling these markets is knowledge of their culture, so that efforts by firms are thoughtful and beneficial to all concerned. This article proposes that the BOP culture is complex and dynamic, characterized by shared aspirations and limited opportunities, a paradox of lacks and abundance, embeddedness in layers of cultural influences, and heterogeneous meanings and groupings in response to externalities. Global firms can use these insights to generate appropriate and effective marketing strategies and tactics that not only open significant new markets, but serve and enhance the lives of those who have been historically marginalized.

ENDNOTES

[1] The term *BOP* is not intended pejoratively. Other terms used more or less synonymously in the literature are bottom of the pyramid, subsistence markets, and low-income consumers

Bibliography

Askegaard, S., Arnould, E.J., and Kjeldgaard, D. (2005) Postassimilationist ethnic consumer research: qualifications and extensions. *Journal of Consumer Research*, **32**, 160–170.

Ger, G. (1997) Human development and humane consumption: well-being beyond the 'good life'. *Journal of Public Policy and Marketing*, **16** (1), 110–125.

Hill, R.P. (2002) Consumer culture and the culture of poverty: implications for marketing theory and practice. *Marketing Theory*, **2** (3), 273–293.

Hofstede, G. (2001) *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*, Sage Publications, Thousand Oaks.

How Did Nokia Succeed in the Indian Mobile Market, While its Rivals Got Hung Up? http://ibef.org/artdisplay.aspx?art_id=16500&cat_id=349&page=1, downloaded February 27, 2009.

Jones, R.K. and Lou, Y. (1999) The culture of poverty and African-American culture: an empirical assessment. *Sociological Perspectives*, **42** (Fall), 439–459.

Leacock, E.B. (1971) Introduction: the culture concept, in *The Culture of Poverty: A Critique* (ed. E.B.Leacock), Simon and Schuster, New York, pp. 7–37.

Lewis, O. (1959) *Five Families: Mexican Case Studies in the Culture of Poverty*, Basic Books, New York.

Lewis, O. (1966) *La Vida*, Random House, New York.

Peñaloza, L. (1994) *Atravesando fronteras/*border crossings: a critical ethnographic exploration of the consumer acculturation of Mexican immigrants. *Journal of Consumer Research*, **21**, 32–54.

Prahalad, C.K. (2005) *The Fortune at the Bottom of the Pyramid*, Wharton School Publishing, Upper Saddle.

Sen, A. (1999) *Development as Freedom*, Anchor Books, New York.

Subrahmanyan, S. and Gomez-Arias, J.T. (2008) Integrated approach to understanding consumer behavior at the bottom of pyramid. *Journal of Consumer Marketing*, **25** (7), 402–412.

Ustuner, T. and Holt, D.B. (2007) Dominated consumer acculturation: the social construction of poor migrant women's consumer identity projects in a Turkish squatter. *Journal of Consumer Research*, **34**, 41–56.

Viswanathan, M. (2007) Understanding product and market interactions in subsistence marketplaces: a study in South India. *Advances in International Management: Product and Market Development for Subsistence Marketplaces*, vol. 20, (eds. M. Viswanathan and J. Rosa), Elsevier, Greenwich, pp. 21–56.

Viswanathan, M., Rosa, J.A., and Harris, J.E. (2005) Decision making and coping of functionally illiterate consumers and some implications for marketing management. *Journal of Marketing*, **69**, 15–31.

global account management: the rationale and motivation

Michael Harvey and Miriam Moeller

Introduction

The coordination of customer management across national boundaries, commonly referred to as *global account management* (GAM), has rapidly emerged as one of the most prevalent corporate responses to the globalization of markets. Yip and Madsen (1996) observe that the need for GAM is driven by the emergence of global customers, global channels (*see* INTERNATIONAL MARKETING CHANNELS), high product development costs, fast-changing technology, and globalized competitors, to name a few. It was hypothesized by Birkinshaw, Toulan, and Arnold (2000) that GAM structures allow the multinational corporation (MNC)/global organization (i.e., the supplier) to increase its information-processing capacity as well as its bargaining power with regard to the global customer. The National Account Management Association has described GAM as a well-articulated management strategy and a "critical competitive weapon" and has stated that GAM "clearly represents the new frontier in Strategic Account Management" (Birkinshaw, Toulan, and Arnold, 2000). The motivation for forming and utilizing a GAM strategy varies widely: it can be due to both positive as well as negative factors (see Table 1).

Strategically, GAM teams serve to ensure that organizational components are configured to best serve global customers effectively/efficiently. In essence, GAM is a means of coordinating activities to address the needs of a single customer in multiple environmental contexts (i.e., countries, markets, etc.) and to ensure the successful coordination and management of these clients globally (Birkinshaw, Toulan, and Arnold, 2000). Thus, the responsibilities of GAM are as follows: (i) creating consistency in the application of policies throughout the world; (ii) coordination of marketing/selling activities to increase sales volume; (iii) effective utilization of marketing strategies and programs in multiple locations; (iv) efficiency of management, in that, there is a central contact of key accounts; (v) establishing a control mechanism relative to key accounts to reduce the probability of account turnover; (vi) improve the two-way flow of communications with key accounts, thereby increasing the knowledge base to improve the quality of goods/services to these global clients; and (viii) awareness as a means to preempt local/global competitors from securing business from these critical customers (Birkinshaw, Toulan, and Arnold, 2000).

The Rationale for Formation of GAM Teams

The rationale for using GAM teams can be derived from a number of perspectives. First, GAM teams provide the parent organization with flexibility that other organizational configurations do not present. That being so, the ability to readily adapt to an increasingly complex global context is important for organizations as they rapidly expand their operations overseas or in configuring their strategic orientation to become more global (Harvey, Novicevic, and Garrison, 2004). Second, the team concept (i.e., being dedicated to one customer globally) increases the level of attention/focus on key decision variables of the global customer. Stated differently, the responsibility for the key account globally is that of the GAM team. This attention to consistency and coordination should alleviate internal decisions that are at cross purposes. (Martins, Gilson, and Maynard, 2004). Third, the GAM team has the ability to modify and/or change its composition (i.e., diversity of team members) to meet changes in the strategic thrust of their customer. As the client of the GAM team modifies its strategic direction (e.g., geographically, product line, or the like) the team can quickly match the direction of the GAM customer (Yip and Madsen, 1996). Thus the rationale used to justify GAM teams is that they can be disbanded/added to more easily than traditional organization structures. Given that the motivation for forming a GAM frequently is based on the expectations of a global client, the ability to shrink the size of the GAM team, as well as to redirect its strategic focus, is a critical element for deciding on using GAM teams.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Motivation for global account management strategy.

	Proactive motivation	
	• Increased sales of existing products/services • Facilitating introduction of new products and services • Coordination of selling/marketing efforts globally • Increased profit due to sales increase and cost containment • Means to reduce competition to supply global accounts and overseas markets	• Building long-term relationships with key customers • Increasing customer dependency • Development of cooperative/synergistic relationships with key customers to ensure global competitiveness • Maintaining strategic consistency, thereby, increasing strategic "fit" with key customers • Global utilization of successful marketing and sales programs • Increased network efficiency and effectiveness globally
Short-run perspective		**Long-run perspective**
	• Pressures from customers to maintain global consistency of products and services • Pressures from customers to reduce prices globally • Concern relating to local competitors in certain regions getting the customer's business • Major concern related to loss of key accounts if GAM structure is not implemented	• Means to effect power shift to key global customers • Dependence on key customers for global sales growth • Need to provide value-added services without commensurate price enhancement • Liability to compete for global customers without GAM • Customers are demanding GAM, but economic and effectiveness justification cannot be supported
	Reactive motivation	

Adapted from: Harvey, M., Myers, M., & Novicevic, M. (2002). Managerial issues associated with Global Account Management, Journal of Management Development, 22(2), 103-129.

The fourth rationale is that the GAM team can distribute the decision making to the point closest to the consumer/environment context. Market-based decision making allows the GAM team to push decision making to the point where the environmental context can be taken into strategic consideration. This allows the GAM team to think globally but to act locally, a key consideration in effectively addressing global clients' needs (Harvey, Myers, and Novicevic, 2002). An example that illustrates this point is the AT&T GAM that serves other global

customers (i.e., those who require worldwide services) by providing them with a worldwide plan as opposed to a collection of individual, country-specific plans. This is not to say that the services offered by AT&T may not be differentiated, as services certainly are to some extent adapted to local environments (e.g., thinking globally but enacting local strategies). From AT&T's perspective, the challenge lies in reconfiguring the way of thinking about their customers' business and how to effectively conduct business in a global context. This modification in strategic thinking incorporates the following four factors: (i) organizational structure, (ii) management processes, (iii) personnel, and (iv) organization culture.

A fifth justification for the formation of GAM teams is to make timely global decisions. More traditional, centralized organization structures are frequently slow in making timely competitive decisions in far-flung global markets. This rationale for GAM addresses the "economies of scope" dimension of clients of the GAM. The more geographically dispersed the clients, the greater the need to move decision makers closer to the foreign markets where their decisions are being made (Brass *et al.*, 2004). By breaking down the central decision-making hierarchy (e.g., utilizing teams) into functional heterarchies, the time responsiveness for GAM clients should be enhanced (Harvey, Myers, and Novicevic, 2002).

Building Trust and Social Capital Through GAM Teams

As the global environment becomes more competitive, strategic HR managers increasingly have to address their new competitive landscape (Bettis and Hitt, 1995). This new landscape is replete with technological advancements leading to the rise of global markets and increase in global competition. The complexity and dynamism of the new economy urges organizations to find innovative ways to cultivate and manage trust and the residual social capital. Thus, the growing importance of strategic global management illuminates the need for the assessment of "capital" within the organization to allow it to maintain its focus on core competencies and viability within the global arena. To stay innovative and ahead of competition, HR must pursue a strategic vision addressing this dynamic and complex issue by tapping those "capital" resources that are vital in maintaining or creating a competitive advantage. One of these resources is constructing social networks to facilitate the development, growth, and management of social capital.

Putnam (1995) defines *social capital* as "the features of social organization, such as networks, norms, and trust that facilitate coordination and cooperation for mutual benefit" (p. 35). Inherent in this definition is the trusting relationship between network sources (Putnam, 1995). Because of the trust component, researchers believe that social capital encourages cooperative behavior (Fukuyama, 1995) and that social capital reduces transaction costs, because trust decreases feelings of suspicion and the likelihood for opportunistic behaviors (Putnam, 1995).

Global organizations possess numerous resources, but perhaps it is their knowledge base that is central to their competitive edge (Barney, 1991). One avenue strategic HR managers are pursuing in the quest of revolutionary knowledge is through the use of GAM teams and the next generation of global virtual teams (GVTs). GAM teams are rapidly being integrated into global organizations, because they are thought to increase speed and efficiency in the creation and transfer of knowledge (Kogut and Zander, 1996). More specifically, GAM teams are thought to enhance the conception and delivery of tacit knowledge since GVTs are characterized by their temporary existence coupled with the potential of their members to multitask simultaneously on multiple projects (Townsend, DeMarie, and Hendrickson, 1998). Because the global environment is in constant flux, it is imperative for GAM team members to develop social networks to facilitate the exchange of thoughts and ideas for the creation of new knowledge with the expectation of creating a competitive advantage. As a result of continued knowledge sharing, the current and future supplier/customer duos will benefit from this synergistic cooperation by a decrease in the learning curve, and by allowing for lower operation costs.

The relationship between organizational knowledge and its competitive advantage lies

in its ability to integrate knowledge utilized in GAM teams (Nahapiet and Ghoshal, 1998). Knowledge is transferred through the GAM team members' relationships with others outside the organization (i.e., social capital). Through the development of social capital by the GAM team members (i.e., social networks), GAM can (i) gain access to financial resources (the GAM will be adequately funded to ensure project completion); (ii) augment the GAM's cultural capital (through collaboration efforts with members of varying functional and cultural expertise); and (iii) increase their human capital (by acting upon the accessibility and timely availability of novel information, the GAM members will increase their "stock" within the organization).

Additionally, those with greater levels of social capital have higher levels of human capital, because they are better positioned to identify and act upon timely information, thereby increasing their chances for more rewarding opportunities (Burt, 1997). The premise behind social capital is that it is an opportunity (Burt, 1997) created by network ties that provide access to diverse resources and information benefits. A result of such information-rich networks is the increasing opportunities for the availability, development, and dissemination of new knowledge. The benefit of such social networks for GAM team lies in the accessibility of information with a reduction in the amount of time and effort required in obtaining nonredundant information.

Social network studies have demonstrated the importance of strong network ties in transferring tacit knowledge across organizational boundaries. These social networks are developed or augmented via participation on team projects and made stronger through frequent team interaction (Bandura, 1986). The information benefits the GAM team can experience are in the access and timing of the information (Burt, 1997). Timely information provides the GAM team, and ultimately the organization, with a first-mover's advantage. Additionally, the structure of the network indicates the amount of nonredundant information that flows from sender to receiver (Nahapiet and Ghoshal, 1998). GAM teams with "rich" social networks (i.e., those made up of functionally and culturally diverse members) are provided more rewarding opportunities to develop and deliver innovative ideas and products. However, social capital is a resource limited by the network (i.e., hierarchy and density) and by the membership that forms the network (Putnam, 1995).

Social capital theory proposes that the networks of relationships create a valuable resource for the creation of knowledge. Additionally, social capital resides within these social networks, resulting from feelings of gratitude and respect (Bourdieu, 1985). Nahapiet and Ghoshal (1998) proposed that it is the web of contacts or connections that creates the pattern of linkages, which create the foundation for social capital. Thus these complex relationships create a process in which social capital is developed and sustained through the process of knowledge transfer. Consequently, the organizations that make use of their collective expertise and knowledge tend to be more innovative and efficient, and it appears that it is the creation and strengthening of social relations that is critical to effectively transfer tacit knowledge throughout the global organization (Hansen, Podolny, and Pfeffer, 2001).

Inherent in the development of GAM teams, as with conventional teams, is the need for trust, since trust has been linked to team performance (O'Hara-Devereaux and Johansen, 1994) and may be substituted as a controlling mechanism in a global virtual context. *Trust*, another component of interorganizational GAM relationships defined by Davis *et al.* (2000), is "the willingness of an individual to be vulnerable to the actions of another person based on the expectation that the other will perform a particular action important to the individual without regard to the ability to monitor or control that other person" (p. 712). Despite the inherent importance of trusting relationships, the strength of the dimension is still dependent upon the type of relationship between the trustor and trustee. The pursuit of trust is therefore a vital component of achieving the desired state of cooperative and synergistic behaviors organizations strive for in the formation and development of GAM teams.

Summary

GAM is an organizational configuration resulting from the rapid expansion of firms into the global marketplace. The GAM team essentially provides a "bridge" for the MNC to evolve into a fully articulated global organization. This organizational metamorphosis provides the organization with an interim means to address the needs of its global consumers and at the same time enables the MNC to become global in its scope. The GAM team's success is based upon the social capital and resulting trust that is built with the key global account that the GAM services. In the future, when organizations have more insight and experience with regard to competing in a global context, newer forms of organizational structures will supersede the present GAM organizational configuration.

Bibliography

Bandura, A. (1986) *Social Foundations of Thought and Action: A Social Cognitive Theory*, Prentice Hall, Englewood Cliffs.

Barney, J. (1991) Firm resources and sustained competitive advantage. *Journal of Management*, **17** (1), 99–120.

Bettis, R. and Hitt, M. (1995) The new competitive landscape. *Strategic Management Journal*, **16**, 7–19.

Birkinshaw, J., Toulan, O., and Arnold, D. (2000) Global account management in multinational corporations: theory and evidence. Paper presented at the Meeting of the Academy of Management Conference, Toronto, Canada.

Bourdieu, P. (1985) The social space and the genesis of groups. *Theory and Science*, **14** (6), 723–744.

Brass, D.J., Galaskiewics, J., Greve, H.R., and Tsai, W. (2004) Taking stock of networks and organizations: a multilevel perspective. *Academy of Management Journal*, **47** (6), 795–817.

Burt, R.S. (1997) A note on social capital and network content. *Social Networks*, **19**, 355–373.

Davis, J., Schoorman, D., Mayer, R., and Tan, H. (2000) The trusted general manager and business unit performance: empirical evidence of a competitive advantage. *Strategic Management Journal*, **21** (5), 563–576.

Fukuyama, F. (1995). Social capital and the global economy: a redrawn map of the world. Foreign Affairs (Sep-Oct), pp. 76–85.

Hansen, M., Podolny, J., and Pfeffer, J. (2001) So Many Ties, So Little Time: A Task Contingency Perspective on Corporate Social Capital in Organizations, Division of Research, Harvard Business School, Boston.

Harvey, M., Myers, M., and Novicevic, M. (2002) Managerial issues associated with global account management. *Journal of Management Development*, **22** (2), 103–129.

Harvey, M., Novicevic, M., and Garrison, G. (2004) Challenges to staffing global virtual teams. *Human Resource Management Review*, **14** (3), 237–256.

Kogut, B. and Zander, U. (1996) What firms do? Coordination, identity, and learning. *Organization Science*, **7** (5), 502–518.

Martins, L., Gilson, L., and Maynard, M. (2004) Virtual teams: what do we know and where do we go from here? *Journal of Management*, **30**, 805–835.

Nahapiet, J. and Ghoshal, S. (1998) Social capital, intellectual capital, and the organizational advantage. *Academy of Management Review*, **23** (2), 242–266.

O'Hara-Devereaux, M. and Johansen, R. (1994) *Global Work: Bridging Distance, Culture and Time*, Jossey-Bass Publishers.

Putnam, R. (1995) Bowling alone: America's declining social capital. *Journal of Democracy*, **6**, 65–78.

Townsend, A., DeMarie, S., and Hendrickson, A. (1998) Virtual teams: technology and the workplace of the future. *Academy of Management Executive*, **12** (3), 17–29.

Yip, G.S. and Madsen, T.L. (1996) Global Account Management: the new frontier in relationship marketing. *International Marketing Review*, **13** (3), 24–42.

marketing's corporate responsibility

Nicolas M. Dahan, Jonathan P. Doh, and Terrence R. Guay

Introduction

Corporate responsibility or, often, corporate social responsibility (CSR), is the notion that companies should operate not just on behalf of their stockholders but in the best interest of society and other stakeholders (including workers, suppliers, environmentalists, and communities). This is controversial for several reasons, particularly since it may undermine the long-held view that companies should aim to maximize the wealth of their owners. However, pressure from a variety of actors, including nongovernmental organizations (NGOs), government, and other stakeholders are forcing companies to rethink their role in society, and to seek to provide benefits to both their owners and society more broadly. The view that CSR can divert management's attention from maximizing shareholder value is countered by the fact that many corporate executives now take seriously this alternate view to strategic management. Their reasons range from agreeing with the concept of CSR (a progressive view of the role of business in a postmodern society) to wanting to avoid the negative publicity that may influence customers to take their business elsewhere (a view more in line with preserving shareholder value) to the hope that voluntary actions will preclude more costly government regulation. Evidence that CSR and economic performance are not mutually exclusive, and can even be complementary, bolsters the case that firms can "do well while doing good." Corporate initiatives in CSR can be complicated by the fact that it is not always obvious what is in society's best interest. Likewise, with regard to marketing policies, nonstockholding stakeholders have expressed expectations that may be financially costly in the short-term and hinder business growth. Socially responsible marketing may however be in a firm's long-term interest.

Conceptual Dimensions of CSR

Stakeholder and institutional theories of CSR. The idea that companies should benefit both their owners (i.e., shareholders) and society more broadly affects business in two direct ways. The first is that CSR accepts that companies must respond to the interests of stakeholders, which is a much broader set of interests than stockholders. The idea that managers should maximize the wealth of their company's shareholders is a fundamental tenet in the management and financial literature, and is widely held in practice among US companies. The economist Milton Friedman famously wrote in *Capitalism and Freedom* that, "Few trends could so thoroughly undermine the very foundations of our free society as the acceptance by corporate officials of a social responsibility other than to make as much money for their stockholders as possible" (p.135). Friedman was concerned that CSR would, in effect, lead to the private sector taking on public sector responsibilities, and that managers are ill-equipped to determine which actions will best serve society's interest. Managers, according to the mainstream view, instead should focus on actions that will increase the value of the company and, in turn, the company's stockholders. Governments and individuals can use their tax revenues and personal wealth, respectively, to achieve social goals.

Over the past two decades, stakeholder theory has challenged the primacy of stockholders in management theory and practice. Stakeholder management, stakeholder theory, and other variants of stakeholder analysis have occupied a great deal of managerial research. Edward Freeman argued that business relationships should include all those who may "affect or be affected by" a corporation. Much of the research in stakeholder theory has sought to systematically address the question of which stakeholders deserve or require management attention. Approaches to this question have focused on relationships between organizations and stakeholders based on exchange transactions, power dependencies, legitimacy claims, or other claims. Mitchell, Agle, and Wood (1997) developed a typology for classifying stakeholders. They proposed a theory of

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

stakeholder identification and salience based on managerial assessments of stakeholders' possession of one or more of three relationship attributes: power, legitimacy, and urgency. Managers should respond in priority to the claims of stakeholders who possess all three characteristics. Researchers have attempted to integrate stakeholder theory with other managerial perspectives, particularly theories of governance and agency. The stakeholder concept is more widely implemented in continental Europe, where key interest groups such as labor unions, suppliers, banks, and government officials are involved to varying degrees in corporate decision making due to tradition or legal requirements.

In a related vein, neoinstitutional theory explains CSR as a corporate response to social pressures setting expectations on how firms should structure and operate their businesses (Campbell, 2007). As societies' concerns for social issues (such as equal opportunities, labor conditions, or the environment) have grown in the past three decades, mounting social pressures have forced a lot of firms to reconsider their place in society.

NGOs are a subset of stakeholders that have become increasingly influential in shaping and expressing these social concerns in recent years. NGOs first rose in prominence in the 1980s, when a range of NGOs, including church and community groups, human rights organizations, and other antiapartheid activists, built strong networks and pressed US cities and states to divest their public pension funds of companies doing business in South Africa. This effort, combined with domestic unrest, international governmental pressures, and capital flight, posed a direct, sustained, and ultimately successful challenge to the white minority rule, resulting in the collapse of apartheid. NGOs have grown in number, power, and influence since the 1980s. Their force has been felt in a range of major public policy debates, and NGO activism has been responsible for significant changes in corporate behavior and governance. Because of the interests that many NGOs represent, including the environment, labor, and human rights, these groups are an important component in discussions on CSR. Consequently, the emergence of NGOs seeking to promote more ethical and socially responsible business practices is beginning to cause substantial changes in corporate management, strategy, and governance (Doh and Teegen, 2003). However, Bhattacharya and Korschun noted that so far, "there has been scant attention in the marketing literature on how these groups [nonprofit partners and nongovernmental organizations] function and what effects they have on constituents' perceptions" (2008, p. 114).

Conceptual applications of CSR to marketing. Bartels (1967) offered a visionary theoretical approach of the social responsibilities and ethics of marketing. He detailed a model very much in line with the aforementioned stakeholder and neoinstitutionalist theories, viewing marketing as a socially embedded process, rather than a purely economic activity. He stressed that marketing operates in a social space where marketing officers must address participants' expectations, which become institutionalized into norms of behavior over time.

As the use of marketing became pervasive in the 1960s and 1970s, and irresponsible policies were uncovered (such as Nestlé's infant formula or Ford's Pinto model), social expectations for its proper practice rose. This point is summarized by Lavidge (1970), "as the impact of marketing on society increases, so does the social responsibility of marketing people" (1970, p. 28). Initially, these expectations were particularly shaped by the new consumerist movement (Abratt and Sachs, 1988). This led to a changed outlook on the insertion of marketing in society, attempting to focus marketing on consumer welfare (improving quality of life) in a sustainable manner. This challenge has been taken on by three conceptions of responsible marketing. Historically, the first stream of research has been a systems-based holistic approach coined "macromarketing," viewing the economic system of production, delivery, and consumption as embedded in social, legal, and political systems (Layton and Grossbart, 2006). This is rather similar to the more recent approach of "societal marketing" (Abratt and Sachs, 1988). A barely emergent view is called the "stakeholder marketing perspective," advocating an outside-in approach that addresses the expectations of various constituents rather

than focusing on an inside-out approach that starts with the firm and its bottom line. This new stream also incorporates the latest developments in stakeholder theory, moving away from the hub-and-spoke view where the firm establishes one-on-one relations with various stakeholders, to a more realistic view based on complex networked relations (Bhattacharya and Korschun, 2008).

The extent of corporate commitment to CSR. The implication of the notion of CSR is that companies should include in their decision-making the impact of their actions on a range of issues, including their employees, the ecological environment, the communities in which they operate, political institutions and processes, and poverty and economic development, among other issues. Many companies have responded by undertaking a number of initiatives addressing these social issues.

There is some debate as to whether such efforts reflect an embrace of CSR by companies, or are simply weak attempts to divert the attention of critics at a minimal cost (Vogel, 2005) or acts of cynicism, simply pretending to care in order to boost sales. "Greenwashing," for example, has come to refer to public relations initiatives by companies to present themselves as ecologically minded, when, in fact, little change has taken place. This is why Brønn and Vrioni (2001) foresaw a rise in skepticism in public opinion, as more firms embrace CSR without being truly committed. Other critics argue that the voluntary nature of CSR schemes means that competitor companies are not bound to follow, and may have a cost advantage. Finally, some believe that companies undertake CSR policies to forestall more stringent, and possibly more expensive, government regulations. Nonetheless, there is some evidence that a new generation of managers is rising to positions of influence in companies, and that there is a greater awareness of the societal impact of companies and interest in addressing such concerns.

CSR and performance. An additional variant of research related to the role of the corporation in modern society comprises those efforts designed to develop and test models of corporate social performance, CSR, and corporate social responsiveness. Empirical studies and reviews of these efforts suggest a link between corporate social responsibility/performance and the economic performance of the corporation (Cochran and Wood, 1984; Griffin and Mahon, 1997). Such relationships suggest that incorporating stakeholders into the strategy-making process may generate social benefits that yield higher financial returns.

Contrary to popular belief, companies are not always opposed to implementing CSR policies or the enactment of environmental regulations (Prakash, Krutilla, and Karamanos, 1996). If the costs of implementing environmental agreements are diffuse and the benefits concentrated, they will support such agreements. This insight suggests that when stakeholders are successful in urging the widespread adoption of standards, corporations may support agreements if they are able to share the costs while preserving individual benefits. In fact, there may be competitive advantages to doing so (*see* COMPETITIVE ADVANTAGE). Companies that conform to environmental regulations promulgated in one jurisdiction gain global first mover competitive advantages when other countries later adopt such standards. Hence, companies may willingly adopt international labor or environmental policies if they believe that such codes will eventually become standard in their industry.

Some companies have incorporated CSR so deeply into their mission statements, corporate culture, and products that they use it as the basis of their competitive strategy. They also find that there is a segment of consumers who make buying decisions based on their perceptions of a company's commitment to CSR (*see* ENVIRONMENTAL CONSUMER BEHAVIOR). Companies can thus differentiate themselves from their competitors, as Green Mountain Coffee Roasters, Whole Foods Market, Google, and Timberland have done. Even global companies in otherwise controversial industries sometimes use CSR to distinguish themselves from competitors (*see* INTERNATIONAL RETAILING). After facing intense criticism for not being socially responsible in Nigeria and elsewhere in the 1990s, Royal Dutch Shell began to champion CSR. As such, the company is regarded by many

stakeholders as more committed to alternative energy, minimizing environmental damage, and improving community relations than peers such as ExxonMobil.

In reviewing the emergent influence of CSR, research suggests that firms experience benefits from improving their social, labor, and environmental performance. These benefits can result from competitive advantages that emanate from improved efficiency and performance and also from reputation and goodwill that result from positive perceptions of the corporation. Further, returns may increase for those firms that adopt such standards because they receive protection or insulation from criticisms. Negative images and reputation will fall on those firms that are perceived to be acting contrary to public interest. In addition, as some firms adopt CSR practices, others may follow, generating a dynamic, reflexive process. Hence, there is motivation for companies to seek collaborations and partnerships with stakeholders under certain conditions.

More specifically, with regard to responsible marketing, the same debate over the economic return of these investments exists and points to mixed empirical results (Abratt and Sachs, 1988). Empirical studies have tested the impact of either CSR policies or specific cause-related marketing initiatives on consumer behavior. The lack of consistency in the results may be due to at least three factors. First, negative CSR information about a firm has been found to have a higher (negative) influence on consumers' product purchase intentions than positive CSR information do (Bhattacharya and Korschun, 2008; Sen and Bhattacharya, 2001). In this sense, responsible business may be more about minimizing the risk of revenue losses than maximizing revenues. Second, different stakeholders have different expectations. Thus, while supporting a certain cause may result in increased purchases from one type of constituency, it may have no effect or even a negative effect on others (Sen and Bhattacharya, 2001). For example, Domino's Pizza faced a boycott from the prochoice movement in the 1980s because the founding CEO James Monaghan was a prolife fervent Catholic who had personally supported financially the antiabortion movement. Third, these investments may yield to long-term benefits, which do not appear immediately (Abratt and Sachs, 1988). With this caveat about the existence of mixed results, as is the case for general studies of the impact of CSR on corporate financial performance, the majority of marketing studies do point out to a positive impact of responsible practices on consumer behavior (Brønn and Vrioni, 2001; Sen and Bhattacharya, 2001). This impact is obviously higher when the firm resorts to cause-related marketing, that is, actively communicates its commitment to a social cause to its consumers. Such initiatives can generate positive media coverage, build corporate reputation, motivate employees and ultimately differentiate the firm's product or service from the competition (Brønn and Vrioni, 2001).

CSR in Practice: Challenges and Opportunities

Challenges. The main challenge in trying to meet social expectations is that these expectations vary across stakeholders as well as across countries (due to differing cultures, religions, political and legal systems, levels of economic development: Brønn and Vrioni, 2001). As pointed out by Lavidge (1970), these expectations will also evolve over time: "some practices which today are generally considered acceptable will gradually be viewed as unethical, then immoral, and will eventually be made illegal" (1970, p. 25). This is why in many instances it is not obvious which corporate action is in society's best interest and which makes CSR implementation complicated, if not controversial.

CSR pressure is more prominent in certain industries or among specific companies that provide goods and services deemed undesirable or harmful to the general public. Companies in the tobacco industry are an obvious example, as is the oil industry. But even companies with otherwise positive contributions to CSR can face considerable criticism when one part of their business comes under scrutiny. Since the 1970s, consumer groups in the United States and Europe have urged the boycott of products sold by Nestlé because the Swiss company markets its infant formula to mothers in poor countries as a substitute for breast milk. Critics contend

that such practices are not socially responsible, since the formula is often mixed with contaminated water, mothers are ill-informed about the relative benefits of breast milk, and scarce financial resources are spent unnecessarily by some of the world's poor. Although Nestlé eventually changed some of its marketing practices to abide by an international code of conduct on breast milk substitutes, allegations that the company was providing health facilities with infant formula at prices below the cost of production instigated another round of boycotts (*see* ETHICAL MARKETING AND MARKETING STRATEGY).

Nike learned this lesson in the 1990s when the company was accused of manufacturing footwear and other apparel in sweatshop conditions, mainly in Asia and Central America. One of the key complicating factors was that Nike did not manufacture anything – it outsourced the production process to other companies which, in turn, created a complex network of suppliers (*see* GLOBAL SOURCING STRATEGY: AN EVOLUTION). Instead, Nike focused its attention on marketing its products to consumers. At first, Nike countered that the sweatshop issue was not the company's problem, since third party companies employed the workers who made Nike products. The criticisms grew more intense, and pressure from the media, consumer boycotts, and labor-oriented NGOs forced the company to develop a set of rules that its suppliers were required to abide by. These included prohibiting child labor, reducing harassment, and improving basic working conditions. While this may have satisfied some of Nike's critics, the actions produced ambiguous consequences. Children who might otherwise have worked in factories turned to other, more hazardous, occupations. Some women who had sought employment to enhance their independence and access to financial resources lost their jobs. Labor issues like this force companies to wrestle with the question of what it means to be socially responsible. From the perspective of Western countries in Europe and North America, labor conditions in developing countries may appear demeaning and exploitative. But from the perspective of emerging markets, the jobs provided by multinational corporations (MNCs) provide opportunities for economic development (*see* EMERGING MARKETS). Companies seeking to be socially responsible need to think hard about the consequences of their CSR strategies.

For some companies, operating in countries where governments are authoritarian, repressive, and corrupt is unavoidable. This is especially true for the natural resources and extractive industries. It is common for such companies to be criticized for cooperating with repugnant regimes, and even strengthening them through the tax revenues provided by foreign investment. The oil company Unocal has faced intense criticism for building a pipeline in Myanmar (formerly Burma), allegedly with forced labor provided by the country's military junta. Other MNCs have withdrawn from Myanmar as a result of pressure by governments, consumers, and shareholders in the United States and Europe, who argue that companies have a social responsibility to not support a brutal regime. Similarly, the Canadian energy company Talisman was forced to sell its stake in an oil project in Sudan, after human rights groups accused the company of providing the government with the necessary revenues to wage a civil war, and major shareholders threatened to sell their shares. Trade sanctions present a related kind of CSR dilemma for companies, since they may be required to withdraw from markets that their home or other governments disapprove of, usually for political reasons (*see* EMBARGOES AND SANCTIONS).

Opportunities. Environmental sustainability, or meeting the needs of the present population without compromising the ability of future generations to meet their own needs, is an increasingly popular concept in the CSR arena. Companies across every industry are urged to reduce their environmental impact, "carbon footprint," and use and disposal of resources. So, for example, companies in the extractive industries like mining and oil are encouraged to minimize their environmental impact by changing production methods or repairing damage. Some timber companies plant new trees after cutting existing ones. A small but growing number of manufacturing companies utilize a closed-loop supply chain, which aims

to reuse or recycle materials throughout the production process, and to take back products when they have reached the end of their useful life (*see* SUPPLY CHAIN MANAGEMENT STRATEGY). Still other companies seek to reduce their environmental impact and energy use in factory and office design. It is fair to say that all of these decisions are a combination of CSR efforts, as well as economic and competitive pressures.

Codes of conduct are a tool that some companies utilize to burnish their CSR credentials (Doh and Guay, 2004). These are voluntary initiatives that companies agree to abide by. Since the 1990s, concerns over the downsides of globalization, particularly in the environmental and labor areas, have contributed to the growth of codes of conduct. Such codes are viewed by some NGOs and other stakeholders as a means to reign in the excesses of globalization. Codes may take the form of international agreements, or be sponsored by international organizations, the private sector, or nonprofit organizations. Rugmark, for example, is a partnership among development and human rights NGOs, companies exporting carpets from India, the Indo-German Export Promotion Council, and UNICEF India to devise and regulate a special label for hand-knotted carpets made without the use of child labor. Similar codes have been adopted by the apparel and other industries with regard to labor issues. Environment-related codes are also increasing in number and scope. The Forest Stewardship Council (FSC) monitors forestry practices and develops a global program combining public awareness, business collaboration, and green marketing. In the early 1990s, the World Wide Fund for Nature and several other major NGOs began exploring the possibility of setting up a voluntary global certification and accreditation system that could be used to verify whether or not wood products were harvested in a socially and environmentally sound manner. Proponents of forest certification sought to show that forests can be managed in a manner that is viable economically without compromising conservation. Through these means, the FSC convinced Home Depot, Wicke's, Lowe's, and Ikea to sell FSC-certified products. The FSC approach is somewhat unique because it relies on institutionalized cooperation between NGOs and businesses, but environmentally oriented NGOs were pivotal in persuading retailers to apply pressure upstream on forestry companies. One of the most prominent codes of conduct is the United Nations Global Compact. The initiative supports businesses in their efforts to align their operations and strategies with 10 universally accepted principles in the areas of human rights, labor, environment, and anticorruption. To date, Global Compact counts over 5100 corporate participants and stakeholders from over 130 countries.

CONCLUSION

Managers' attention to CSR is unlikely to diminish any time soon. Globalization is increasing the web of linkages between companies, governments, consumers, media, NGOs, activists, and other groups around the world. It has become almost impossible for a company's actions in one corner of the world to go unnoticed elsewhere, particularly in richer countries where most MNCs – and their critics – are headquartered. Companies essentially have two choices in addressing CSR in their marketing and overall business strategies. The first is to react to crises as they develop, as Nike, Shell, and Nestlé initially did. The second and, in all likelihood, more successful option is to be proactive in anticipating which dimensions of their business operations are likely to generate controversy, developing responses in cooperation with critics, and integrating CSR into the company's business model.

Bibliography

Abratt, A. and Sachs, D. (1988) The marketing challenge: towards being profitable and socially responsible. *Journal of Business Ethics*, **7/7**, 497–507.

Bartels, R. (1967) A model for ethics in marketing. *Journal of Marketing*, **31**, 20–26.

Bhattacharya, C.B. and Korschun, D. (2008) Stakeholder marketing: beyond the Four Ps and the customer. *Journal of Public Policy and Marketing*, **27/1**, 113–116.

Brønn, P.S. and Vrioni, A.B. (2001) Corporate social responsibility and cause-related marketing: an overview. *International Journal of Advertising*, **20**, 207–222.

Campbell, J.L. (2007) Why would corporations behave in socially responsible ways? An institutional theory of

corporate social responsibility. *Academy of Management Review*, **32/3**, 946–967.

Cochran, P.L. and Wood, R.A. (1984) Corporate social responsibility and financial performance. *Academy of Management Journal*, **27**, 42–56.

Doh, J.P. and Guay, T.R. (2004) Globalization and corporate social responsibility: how nongovernmental organizations influence labor and environmental codes of conduct. *Management International Review*, **44/3**, 7–30.

Doh, J.P. and Teegen, H. (2003) *Globalization and NGOs: Transforming Business, Governments, and Society*, Praeger, Westport.

Griffin, J.J. and Mahon, J.F. (1997) The corporate social performance and corporate financial performance debate: twenty five years of incomparable research. *Business and Society*, **36/1**, 5–15.

Lavidge, R.J. (1970) The growing responsibilities of marketing. *Journal of Marketing*, **34/1**, 25–28.

Layton, R.A. and Grossbart, S. (2006) Macromarketing: past, present, and possible future. *Journal of Macromarketing*, **26**, 193–213.

Mitchell, R.K., Agle, B.R., and Wood, D.J. (1997) Toward a theory of stakeholder identification and salience: defining the principle of who and what really counts. *Academy of Management Review*, **22**, 853–886.

Prakash, A., Krutilla, K., and Karamanos, P. (1996) Multinational corporations and international environmental policy. *Business and the Contemporary World*, **8** (3/4), 119–144.

Sen, S. and Bhattacharya, C.B. (2001) Does doing good always lead to doing better? Consumer reactions to corporate social responsibility. *Journal of Marketing Research*, **38**, 225–243.

Vogel, D. (2005) *The Market For Virtue: The Potential And Limits Of Corporate Social Responsibility*, Brookings, Washington, DC.

family conglomerates

Daekwan Kim and S. Tamer Cavusgil

Family conglomerates are diversified family-owned businesses that are quite common in numerous emerging markets including India, Indonesia, Korea, Mexico, Philippines, Taiwan, Thailand, and Turkey, and many other countries (Kim, Kandemir and Cavusgil, 2003). They are important not only because of their role in their respective economies but also because of their potential as local business partners for multinational companies. Indeed, they are part of the economic landscape in most emerging markets. They are known as *Chaebols* in Korea, *Business Houses* in India, *Holding Companies* in Turkey, *Bumiputra* in Malaysia, and *Groupos* in Latin America. Family conglomerates are one of the most prevalent business structures with significant roles in these markets.

Owing to their scale and ownership nature, family conglomerates exhibited considerable resilience over the years with their ownership transcending several generations. Their strong presence in their respective economies implies much political clout in governmental relations, and recognition with customers. Interestingly, many family conglomerates seem to have successfully transformed themselves over time in order to sustain their competitiveness. Many are playing a crucial role in introducing new technologies and new products (*see* LAUNCH STRATEGIES; NEW-PRODUCT FORECASTING; SUCCESS FACTORS FOR NEW-PRODUCT DEVELOPMENT) to their respective economies.

Despite variations in organizational and national cultures (*see* SOCIETY, CULTURE, AND GLOBAL CONSUMER CULTURE) surrounding them, these family conglomerates from different countries tend to evolve in much the same way. Quite often their origins and growth can be attributed to the special relationship they have with the government and the fast growing economy (Cavusgil, Kandemir and Kim, 2003). Market expansion such as diversification and internationalization are commonly adopted growth strategies. Ironically, family conglomerates both benefited from and had to adapt to their business environments.

Regarding the sustainability of family conglomerates in emerging markets, the optimistic view is that family influence within family conglomerates will remain strong due to accumulated wealth and stockholder power. Although the ownership of family conglomerates will be dispersed so that one family is less able to control a conglomerate, it is highly likely that a family as a strong collective owner can control a conglomerate utilizing the majority ownership gained through the cross investment. While some subsidiaries of family conglomerates in emerging markets are becoming publicly held, the founder's family still has significant influence and control on the management. Furthermore, as the second or third generation inherits the business, more family conglomerates hire professional managers, leaving family members a less direct role. This may help many family conglomerates solidify their role and presence in the respective markets.

The pessimists believe that family conglomerates will face pressure from government as well as domestic and international communities to reduce the number of core businesses in order to improve efficiency. Potential leadership vacuum could be another issue as the second or third generation takes the leadership role (Church, 1993). In addition, family conglomerates often reveal some degree of inefficiency due to the ownership structure (Khanna and Palepu, 1997; Schulze *et al.*, 2001). Allegiance to owners is a characteristic of family conglomerates (Drozdow and Carroll, 1997; Hayashibara, 1997) but not always good for their business efficiency. Therefore, it is possible that global pressures accentuate those inefficiencies and reduce family conglomerates to a few small businesses in the future.

Given their unique *competencies* (*see* CORE COMPETENCIES) in their respective economies, it is critical for multinational companies contemplating business in such markets to understand their role. As potential local *competitors* (*see* COMPETITOR ANALYSIS), family conglomerates will be formidable barriers for most multinational firms. At the same time, family conglomerates can take much of the hassle and difficulty out of entering and succeeding in emerging markets as collaborators and business partners. Indeed, family conglomerates have

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

much to offer to multinational companies including a local business network, government contacts, consumer recognition, and established *channels* (*see* INTERNATIONAL MARKETING CHANNELS), among others. In turn, family conglomerates also stand to gain from the new business opportunities and know-how that multinational companies offer. By identifying the complementarities that family conglomerates offer, multinational companies can enjoy a more frictionless entry into emerging markets.

Bibliography

Cavusgil, S.T., Kandemir, D. and Kim, D. (2003) The drivers for the evolution of family conglomerates in emerging markets *Bogazici Journal*, **17** (1), 23–44.

Church, R. (1993) The family firm in industrial capitalism: international perspectives in hypotheses and history. *Business History*, **35** (4), 17–43.

Drozdow, N. and Carroll, V.P. (1997) Tools for strategy development in family firms. *Sloan Management Review*, **39** (1), 75–88.

Hayashibara, M. (1997) From family business to multinational. *Asian Journal*, **33** (11), 19.

Khanna, T. and Palepu, K. (1997) Why focused strategies may be wrong for emerging markets. *Harvard Business Review*, **75** (4), 41–51.

Kim, D., Kandemir, D. and Cavusgil, S.T. (2003) The role of family conglomerates in emerging markets: what Western companies should know. *Thunderbird International Business Review*, **46** (1), 13–38.

Schulze, W.S., Lubatkin, M.H., Dino, R.N. and Buchholtz, A.K. (2001) Agency relationships in family firms: theory and evidence. *Organization Science*, **12** (2), 99–116.

international franchising

Barry Quinn

The Franchise Method

Franchising has become a major driving force in the globalization of service businesses. Changes in the global trading environment such as the lifting of restrictive legislation and a greater homogeneity in buyer behavior have resulted in the further spread of franchising activity from the developed economies toward emerging markets in Asia, South America, and Europe. While traditionally associated with the fast food restaurant sector, franchising has been increasingly employed by companies in a range of sectors, including retailing, hotels, car hire, and industrial services. The term *franchising* has been used to describe a wide variety of business activities, but the contemporary franchise system commonly in use is known as *business format franchising*. There are clear advantages to adopting franchising as an international expansion vehicle and these are linked to the relatively low financial resources, lower risk, and the local knowledge offered by the franchisee.

International Expansion

The United States has traditionally been the world's single largest franchise market. However, franchising is now used widely by US and non-US companies as a strategy for growth in both developed economies and emerging markets. Market saturation has been cited as a key factor for franchisors in the United States, Canada, Western Europe, and Japan (Hoffman and Preble, 2004), and international expansion has generally occurred in markets geographically and *culturally close* (*see* MARKETING ASPECTS OF PSYCHIC DISTANCE; MARKETING ASPECTS OF CULTURAL DISTANCE) to the home market. For instance, Canada has been the preferred first destination for US franchisors (see, for instance, Walker and Etzel, 1973; Hackett, 1976; Walker and Cross, 1989). Evidence has been found to support the notion that franchising organizations only make international moves after reaching a significant level in their domestic operations (Walker, 1989; Aydin and Kacker, 1990; Welch, 1990). An initial domestic presence is viewed as crucial for successful international expansion in that it aids the learning process that will be useful for later international expansion and that it creates a widespread network that, by itself, becomes a very tangible statement to potential franchisees, both local and foreign. However, an exception to this "rule" can be seen in the case of *retailing* companies (*see* INTERNATIONAL RETAILING). As service sector companies have developed franchise operations in their respective domestic markets, some retail companies only employ franchising as a strategy in international markets. In other words, franchising is adopted as an alternative strategy to the core operating presence developed within the domestic market (Quinn and Alexander, 2002).

Emerging Market Opportunities

The fundamentally expansionist nature of franchising would suggest that foreign markets will be perceived as providing favorable opportunities for growth, regardless of the level of development of the domestic market (see, for instance, Hackett, 1976; Hopkins, 1996). It is increasingly the case that franchisors are looking for opportunities in emerging markets (Welsh *et al.*, 2006) and franchising is being seen as one way for nations to grow their economies. This is the case in Asian markets such as China and Singapore (Choo *et al.*, 2007; Wang *et al.*, 2008). Such markets offer clear opportunities to franchisors, including high growth and demand for western products and services, urban populations, rising middle classes, and attractive legislative environments. Franchising is of great benefit to transitional economies, particularly the former communist economies of Eastern Europe where entrepreneurship and business development skills have been lacking (Welsh and Swerdlow, 1991; Anttonen *et al.*, 2005).

Forms of Activity

Master/area franchising, joint venture franchising, direct investment, and direct franchising are the major forms of franchising a firm can choose from when it decides to enter a foreign

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

market by the franchising route. In master/area franchising agreements, such as those favored by the Body Shop and the convenience store retailer 7-Eleven, the franchisor grants the master franchisee the right to subfranchise the franchisor's concept to others within an exclusive territory, creating a tripartite franchise relationship. For the franchisor, such arrangements mean that most of the work involved in expanding the operation in the foreign market is carried out by the foreign franchisee, thereby reducing the demands on the franchisor. However, control of the quality of the network's operations is crucial and difficult to maintain (Quinn and Doherty, 2000). The international franchisor will seek to maintain quality and standards through provision of adequate support, in terms of personnel and resources. There is the danger that franchisors may, in practice, underestimate the social, economic, and cultural differences of another country, particularly given the standardization ethos of franchising. This scenario becomes more likely as franchisors seek to avail of the growing opportunities in emerging markets that are generally culturally and geographically distant from the domestic market. Conflict can arise because of the considerable cultural distance (also *see* MARKETING ASPECTS OF CULTURAL DISTANCE) between the two parties, which increases their tendency to see the same situation in quite different ways. Thus, the strength of the relationship between the franchise partners is of paramount importance.

Bibliography

Anttonen, N., Tuunanen, M., and Alon, A. (2005) The international business environments of franchising in Russia. *Academy of Marketing Science Review*, **5**, 1–18.

Aydin, N. and Kacker, M. (1990) International outlook for US-based franchisors. *International Marketing Review*, **7** (2), 43–53.

Choo, S., Mazzarol, T., and Soutar, G. (2007) The selection of international retail franchisees in East Asia. *Asia Pacific Journal of Marketing and Logistics*, **19** (4), 380–397.

Hackett, D.W. (1976) The international expansion of US franchise systems – status and strategies. *Journal of International Business*, **7** (1), 65–75.

Hoffman, R.C. and Preble, J.F. (2004) Global franchising: current status and future challenges. *Journal of Services Marketing*, **18** (2), 101–113.

Hopkins, D.M. (1996) International franchising: standardisation versus adaptation to cultural differences. *Franchising Research: An International Journal*, **1** (1), 15–24.

Quinn, B. and Alexander, N. (2002) International retail franchising: a conceptual framework. *International Journal of Retail and Distribution Management*, **30** (5), 264–276.

Quinn, B. and Doherty, A.M. (2000) Power and control in international retail franchising: evidence from theory and practice. *International Marketing Review*, **17** (4/5), 354–372.

Walker, B.J. (1989) *A Comparison of International vs Domestic Expansion by US Franchise Systems*, International Franchise Association, Washington, DC.

Walker, B.J. and Cross, J. (1989) A progress report on the scope of international expansion by US franchise systems. Proceedings of the Annual Conference of the Society of Franchising, 29–31 January, Bal Harbour.

Walker, B.J. and Etzel, M.J. (1973) The internationalisation of US franchise systems: progress and procedure. *Journal of Marketing*, **37** (2), 38–46.

Wang, Z., Zhu, M., and Terry, A. (2008) The development of franchising in China. *Journal of Marketing Channels*, **15** (2/3), 167–184.

Welch, L.S. (1990) Internationalisation by Australian franchisors. *Asia Pacific Journal of Management*, **7** (2), 101–121.

Welsh, D.H.B., Alon, A., and Falbe, C.M. (2006) An examination of international retail franchising in emerging markets. *Journal of Small Business Management*, **44** (1), 130–149.

Welsh, D.H.B. and Swerdlow, S. (1991) Opportunities and challenges for franchisors in the USSR: preliminary results of a survey of Soviet university students. Proceedings of the International Society of Franchising, University of St. Thomas Institute for Franchise Management, Minneapolis, MN, February 1991.

global sourcing strategy: an evolution

Masaaki Kotabe and Janet Y. Murray

As global competition has accelerated the speed of technological obsolescence for most products, companies can no longer survive simply by adopting a polycentric, country-by-country approach to international business. If companies with a new product do follow a country-by-country approach to enter foreign markets over time, a globally oriented competitor will likely overcome their initial competitive advantages by blanketing the world markets with similar products in a shorter time frame. Increasingly, how to source globally has become a critical strategic decision that is influenced by the capabilities needed to compete.

Without established sourcing plans, distribution, and service networks, it is extremely difficult to simultaneously exploit both emerging technology and potential markets worldwide. The increased pace of new product introduction and reduction in innovational lead time calls for more proactive management of locational and corporate resources on a global basis. In this article, we emphasize the choices companies make to perform activities either inside the firm or have those activities performed by others, anywhere in the world – which we call *global sourcing strategy*. Global sourcing strategy, therefore, refers to the management of (i) logistics (*see* DESIGNING A GLOBAL SUPPLY CHAIN: OPPORTUNITIES AND CHALLENGES) identifying which production units will serve which particular markets and how components will be supplied for production and (ii) the interfaces among R&D (*see* GLOBAL PRODUCT R&D), manufacturing/operations, and marketing (*see* GLOBAL MARKETING STRATEGY; MARKETING STRATEGY IMPLEMENTATION) on a global basis. Global sourcing strategy requires a close coordination among R&D, manufacturing/operations, and marketing activities across national boundaries (Kotabe, 1992).

Global Sourcing Phenomenon

In a hypercompetitive and uncertain global business environment (*see* SOCIETY, CULTURE, AND GLOBAL CONSUMER CULTURE) coupled with a more even distribution of supply capabilities worldwide, an increasing number of large and small firms either produce in lower-cost locations or outsource goods and services from lower-cost producers. To create a sustainable competitive advantage over their rivals, firms realize that it is imperative to continuously create and acquire capabilities. In addition to securing lower costs from global suppliers, firms increasingly outsource to gain access to suppliers' capabilities. Thus, the core driver of the latest form of global outsourcing (i.e., both onshore and offshore) is the heightened organizational and technological capacity of firms in decoupling and coordinating a network of remotely located external suppliers performing an intricate set of activities. Hence, how to source globally has become a critical strategic decision that is influenced by the capabilities needed to compete and help sustain a firm's competitive advantage.

Although firms have embraced global sourcing of goods and services, they have experienced mixed results. Gottfredson, Puryear, and Phillips (2005) found that about 50% of firms in their sample reported that their outsourcing programs fell short of expectations. Only 10% were highly satisfied with the cost savings, and 6% were highly satisfied with their offshore outsourcing overall. Other researchers (Leiblein, Reuer, and Dalsace, 2002) have even suggested that outsourcing may not be related to performance. Owing to the inconclusive performance outcomes, practitioners have started to question whether universally prescribing global outsourcing is the right way to go.

One plausible argument is that based on a "balance" perspective, there is an optimal degree of outsourcing. The outsourcing–performance relationship takes on an inverted-U shape, implying that as firms deviate further from their optimal degree of outsourcing, by either outsourcing (or insourcing) and offshoring (or onshoring) too much, their performance will suffer disproportionately. So, the key question for sourcing firms is how much global sourcing they should engage in, to achieve desirable performance.

Another plausible argument for the inconclusive sourcing performance findings is that

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Different sourcing strategies.

Ownership Aspect	*Locational Aspect*	
	Domestic Sourcing	*Foreign Sourcing*
Insourcing (intrafirm sourcing)	Onshore insourcing	Offshore insourcing
Outsourcing (contractual sourcing)	Onshore outsourcing	Offshore outsourcing

desirable sourcing performance necessitates the sourcing strategy to achieve a strategic "fit" with the environment. Indeed, researchers have theorized that the appropriateness of a particular strategy is based on its "coalignment" or "fit" with environmental contingencies (Drazin and Van de Ven, 1985). Using contingency theory to examine the environment–strategy coalignment effect on performance, we believe that the environment and strategy interact in a dynamic process, and that a match between them would exert a positive impact on performance. Thus, firms that can adapt their global sourcing strategy effectively to both internal and external factors are likely to achieve better performance.

We focus on global sourcing as it adds many more complexities that do not apply to domestic sourcing strategy. In developing viable global sourcing strategies, firms must consider not only manufacturing and delivery costs, the costs of various resources, and exchange rate fluctuations, but also the availability of infrastructure (including transportation, communications, and energy), industrial and cultural environments, the ease of working with foreign host governments, and other factors. Furthermore, the complex nature of global sourcing strategy spawns many barriers to its successful execution. In particular, logistics, inventory management, distance, nationalism, and a lack of working knowledge about foreign business practices, are some of the major operational problems encountered by both United States and foreign multinational firms engaging in global sourcing.

Intuitive arguments, like "focusing on core competency" and "strategic sourcing," are often made to legitimize the trends toward more global outsourcing. We first discuss the recent trends in global sourcing strategy. Then, we highlight the advantages and disadvantages of global sourcing, by providing a list of intuitive arguments for each. We then attempt to explain global sourcing levels and how these relate to performance based on the two complementary perspectives of "balance" and "fit." By synthesizing these two perspectives, we introduce existing theories of sourcing in this article.

Trends in global sourcing. The primary objective of global sourcing strategy is for the firm to exploit both its own and its suppliers' competitive advantages and the comparative locational advantages of various countries in global competition. From a contractual point of view, the global sourcing of intermediate products such as components and services by firms takes place in two ways: (i) from the parents or their foreign subsidiaries on an "intrafirm" basis (i.e., insourcing) and (ii) from independent suppliers on a "contractual" basis (i.e., outsourcing). Similarly, from a locational point of view, multinational firms can procure goods and services either (i) domestically (i.e., onshoring) or (ii) from abroad (i.e., offshoring) (*see* OFFSHORING AND MARKETING). This leads to a matrix of possible choices presented in Table 1.

In the last two decades, we have witnessed three waves of global sourcing. The first wave, starting in the mid-1980s, was primarily focused on global sourcing of manufacturing activities. Therefore, research was conducted primarily on manufacturing firms. Large manufacturing firms increasingly set up their operations globally and began to use suppliers from many countries to exploit best-in-world sources (Quinn and Hilmer, 1994). Consequently, supply chains (*see* SUPPLY CHAIN MANAGEMENT STRATEGY) became more global and complex, with manufacturing firms sourcing from suppliers in many countries for raw materials, intermediate, and final products.

A second wave began to occur in the early 1990s, when firms started eliminating their information technology (IT) departments that had

Table 2 Recent waves in global sourcing.

Time Period	*First Wave (since 1980s)*	*Second Wave (since Early 1990s)*	*Third Wave (since Early 2000s)*
Type of activity	Manufacturing	Information technology	Business processes
Destinations	China, Central and Eastern Europe, Mexico, and others	India, Ireland, and others	India, Pakistan, South Africa, and others
Type of firms	Manufacturing	Manufacturing, banks, and others	Financial services, and services, more generally
Primary motives	Reduction in labor costs	Obtaining enough skilled programmers and cost reduction	Reduction in labor costs and round-the-clock service provision

grown substantially. As IT itself had become commoditized and many firms had little interest in developing new information systems in-house, this IT outsourcing wave spawned the growth of specialist providers, such as EDS and Accenture. Global sourcing mostly involved labor-intensive and standardized programming activities, which could be easily sourced from low-cost locations like India. The rise of commercial applications for a wide range of firm activities, epitomized in enterprise resource planning systems, also implied that a marketplace had developed where independent suppliers could make competitive offerings.

A third wave, characterized as the offshoring movement, began in the early 2000s. We have witnessed the rise of business process outsourcing that extends beyond IT services to a range of other services related to accounting, human resource management, finance, sales, and after-sales services such as call centers. It is this third wave of business process outsourcing that has generated so much publicity. Many are concerned that foreign business processes suppliers may be moving up the knowledge chain more rapidly than expected by sourcing firms. Such knowledge transfer could, in the long run, undermine sourcing firms' ability to differentiate themselves from their foreign suppliers. Indeed, such hollowing-out concerns have previously been raised about outsourcing of manufacturing activities (Bettis, Bradley, and Hamel, 1992; Kotabe, 1998). We summarize our argument on these recent waves of global sourcing in Table 2.

Global Sourcing Strategy and Performance

It is widely suggested that global sourcing helps improve performance, particularly cost effectiveness (Trent and Monczka, 2003). Firms located in developed countries often find that labor costs are excessive, compared to the value that is added to their products. At the other extreme, some global sourcing may be driven by knowledge concerns. Some inputs, such as liquid-crystal displays and technical expertise, may be available only in certain other countries, thus making global sourcing not a choice but an imperative. As for the sourcing of many raw materials, domestic sourcing is not an option since many raw materials are unavailable domestically. Certain intermediate products tend to be sourced from locations near the source of raw materials. Another argument in favor of global sourcing is that it enables a firm to produce closer to its customer markets, thereby increasing access to its customers and obtaining critical market knowledge for product development (*see* GLOBAL PRODUCT DEVELOPMENT). For instance, Japanese manufacturing firms have, over time, replicated supply chains in North America and Europe to operate closer to these markets. Production and sourcing experience in these regions has also enabled them to improve their product offerings. Another reason to opt for global sourcing is that demand from various regions can be pooled, thus achieving maximum scale and bargaining power through single sourcing from a foreign supplier.

Table 3 Arguments for and against outsourcing.

The Case for Outsourcing	*The Case Against Outsourcing*
Strategic focus/reduction of assets Through outsourcing activities a firm can reduce its level of asset investment in manufacturing and related areas. Therefore, stock markets usually react favorably to outsourcing since more or less similar absolute profit levels can be obtained with lower fixed investments. Furthermore, outsourcing can help the management of a firm redirect its attention to its core competencies, instead of having to possess and update a wide range of competencies.	Interfaces/economies of scope Firms may benefit from internalizing production through scope economies. Manufacturing firms, in their outsourcing decisions, ought to reflect on the interfaces among R&D, manufacturing, and marketing. If there are important interfaces between activities, decoupling them into separate activities performed by different suppliers will generate less than optimal results.
Strategic flexibility Outsourcing may increase the firm's strategic flexibility. By using outside sources, it is much easier to switch from one supplier to another. If an external shock occurs, firms are able to react quickly by simply increasing or decreasing the volumes obtained from an external supplier. If the same item were produced in-house, the firm would not only incur high restructuring costs but also a much longer response time to external events.	Hollowing out Firms that outsource activities excessively are hollowing out their competitive base. Once activities have been outsourced, it tends to become difficult to differentiate a firm's products on the basis of these activities. Furthermore, a firm could lose bargaining power vis-à-vis its suppliers because its suppliers' capabilities may increase relative to those of the firm.
Avoiding bureaucratic costs Rising production costs are associated with internal production, due to a lack of a price mechanism and economic incentives inside a firm. As a consequence, firm efficiency will suffer.	Opportunistic behavior External suppliers may behave opportunistically as their incentive structure varies widely from that of the outsourcing firm. Opportunistic behavior allows a supplier to extract more rents from the relationship than it would normally do, for example, by supplying a lower than agreed-on product quality or withholding information on changes in production costs.
Relational rent In recent years, many researchers have argued that certain relationships with external suppliers can help create a competitive advantage. By outsourcing items on the basis of idiosyncratic and valuable relationships with suppliers, firms may be able to innovate, learn, and reduce transaction costs.	Rising transaction and coordination costs Excessive outsourcing may lead to high coordination costs. Firms are limited in their capacity to work with outside suppliers as partners and therefore, have to prioritize outside partners. If they simultaneously invested time and attention to all outside suppliers, this would induce very high coordination costs.

(*continued overleaf*)

Table 3 (*Continued*).

The Case for Outsourcing	*The Case Against Outsourcing*
	Limited learning and innovation A form of learning that is deemed especially important for attaining tacit knowledge is learning by doing. The supplier may acquire tacit knowledge by performing the activity; consequently, the outsourcing firm cannot appropriate all benefits. Appropriation of innovation and rents is always a problem in buyer–supplier relationships because both parties will try to obtain as many private benefits as possible. Furthermore, it may become more difficult to innovate, owing to the different incentives available and the subsequent lack of interfaces between firms.

On the other hand, there are disadvantages associated with global sourcing. One major problem is "cultural differences" between buyers and their foreign suppliers (*see* BASE OF THE PYRAMID MARKETS: CULTURE INSIGHTS AND MARKETING IMPLICATIONS). Indeed, differences such as institutional and language problems may affect a relationship negatively. This raises another layer of issues related to the long-term sustainability of firms' core competencies, particularly when firms begin to increase reliance on independent suppliers through outsourcing (for a more extensive discussion of outsourcing and core competencies, see Mol, 2007). There are two opposing views of the long-term implications of outsourcing. One school of thought argues that many successful companies have developed a dynamic organizational network through increasing cross-border joint ventures, subcontracting and licensing activities (Miles and Snow, 1986). This flexible network system, also known as *supply-chain alliances*, allows each participant to pursue its particular competence. Each network participant is complementing rather than competing against the other participants for the common goals. The other school of thought argues that while a firm may gain short-term advantages, there could also be negative long-term consequences. As the firm becomes more reliant on its independent suppliers, it may not be able to keep abreast of constantly evolving design and engineering technologies without engaging in those developmental activities (Kotabe, 1998). Consequently, the firm encounters the inherent difficulty in sustaining its long-term competitive advantages. In other words, over time a firm's technical expertise and capability surplus vis-à-vis its foreign suppliers may diminish to the point that its value added is limited, and it may become more like a trading company. Thus, based on the arguments for and against outsourcing, we need to synthesize our thinking on outsourcing and performance. A summary of these opposing arguments is presented in Table 3.

A "balance" perspective. A "balance" perspective offers insights on the sourcing strategy–performance relationship. The underlying argument of a "balance" perspective is that firms that outsource all of their activities run into a multitude of problems, such as a lack of innovation and bargaining power, and an inability to be distinct in the eyes of the customer. However, firms that only insource fail to use the powerful incentives supplied by markets, thus becoming bureaucratic and inefficient. Therefore, outsourcing some but not

all activities provides the best solution overall, and there is an optimal degree of outsourcing.

We believe a similar line of reasoning can apply to the degree of internationalization of sourcing (i.e., onshoring and offshoring) and how that affects performance. More specifically, there are advantages and disadvantages associated with global sourcing, as we highlighted above. As a firm does more offshoring (particularly, offshore outsourcing), the disadvantages become larger to the point where they severely impede performance. If firms do not use offshoring at all, they cannot enjoy any of the advantages of offshoring, such as having a wider supply base from which to choose. This line of reasoning is consistent with research in international business; it is, for instance, indirectly suggested by Dunning's (1993) treatment of international sourcing, and neoinstitutional economics traditions, particularly the transaction costs framework (Williamson, 1985).

Williamson (1985) distinguishes between production and transaction costs. Production costs refer to the costs of producing a good or a service, and transaction costs represent all the costs incurred as the product moves from one supply-chain partner to the next. When firms use offshore outsourcing by procuring from foreign suppliers, it may help reduce their production costs. In some instances, a local supplier's production costs may be lower than those of foreign suppliers, but this is often the exception and not the rule. Transaction costs, on the other hand, tend to be higher for such offshoring, as there are many types of institutional, cultural, and language barriers that must be overcome.

The cost of searching for supply sources abroad, whether internal or external sources, is somewhat higher than that for local supply sources. The cost of evaluating those foreign supply sources is much higher, as the evaluation costs are strongly related to the familiarity that decision makers have with the other party. Since firms are likely to be less familiar with foreign supply sources and decision makers may not be able to draw on their networks in helping them evaluate these sources, this induces substantial evaluation costs. Rangan (2000) uses this argument to explain why buying firms are much more likely to choose a domestic rather than a foreign supplier, even when the physical distance between the buyer and each of these suppliers is the same.

We argue that offshoring is a balancing act between production and transaction costs. Firms need to find the proper balance between domestic and foreign supply sources (using onshoring and offshoring) if they wish to locate on the top of the curve and obtain the highest possible performance. They can achieve this by using foreign sources for part, but not all of their sourcing. Sourcing everything from abroad produces poor performance results because the disadvantages of offshoring, like the hollowing-out argument, become too large. Focusing all efforts on onshoring, however, is a serious form of myopia with equally disastrous effects on firm performance, primarily because the firm is not capitalizing on important opportunities to improve competitiveness. A graphic illustration of our argument is presented in Figure 1.

The balance perspective is therefore summarized as follows: Some activities are best outsourced globally while others ought to be integrated (from a performance perspective). A firm can enjoy optimal performance when it correctly outsources and integrates all activities. Similarly, the firm also needs to balance between onshoring and offshoring activities. This produces a pattern of an inverted U-shaped (negatively curvilinear) relationship between outsourcing and performance, with the top of the curve presenting the performance optimum.

A "fit" perspective. Despite the heightened publicity of global sourcing, many firms have been highly dissatisfied with their sourcing performance. The problem may be due to the fact that many researchers and practitioners have adopted a deterministic view in evaluating the global sourcing strategy–performance relationship, without exercising caution that such a view tends to overgeneralize the sourcing benefits.

Researchers often adopt the contingency approach in representing a "fit" perspective of the environment–strategy–performance relationship. Extant research has confirmed that some environmental factors indeed exerted moderating effects on the sourcing strategy–performance relationship. In the manufacturing

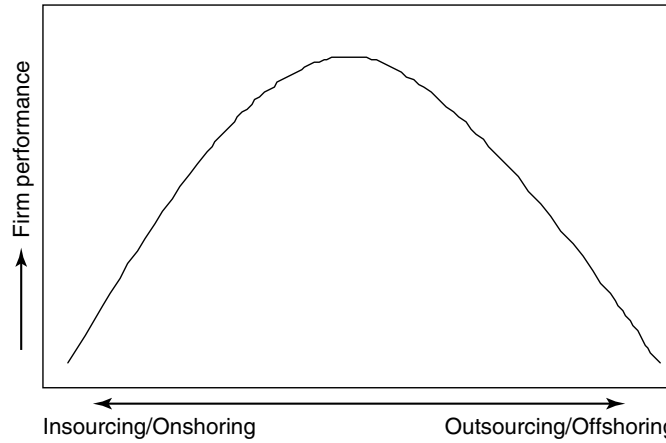


Figure 1 A curvilinear relationship between the degree of global outsourcing and firm performance.

context, Murray, Kotabe, and Wildt (1995) concluded that the financial performance advantage of global insourcing over global outsourcing of nonstandardized (i.e., major) components strengthened with increased product innovations, process innovations, and asset specificity.

Using foreign firms manufacturing in China as subjects of their study, Murray, Kotabe, and Zhou (2005) found that global outsourcing of major components (in the form strategic alliance-based sourcing) did not affect market performance. Instead, product innovativeness and technological uncertainty moderated such a relationship. Specifically, at low levels of product innovativeness/technological uncertainty, the use of strategic alliance-based sourcing of major components by the sourcing firm is positively related to market performance. However, at higher levels of product innovativeness/technological uncertainty, the sourcing–performance relationships become negative.

In refuting the popular arguments that insourcing or outsourcing will lead to superior performance, they found that sourcing strategy per se did not significantly affect performance. Instead, the sourcing strategy–performance relationship was driven by factors underlying sourcing strategy choice. They further cautioned against the universalistic normative implications for firms deciding on whether to insource or outsource their value-chain activities and stressed the value of contingency-based theoretical approaches.

As discussed earlier, global sourcing of services did not take place until the second wave of global sourcing; therefore, extant literature on global sourcing of services (*see* SERVICES MARKETING STRATEGY; SERVICE INNOVATION MANAGEMENT) is limited when compared to that in manufactured goods. Murray and Kotabe 1999 found that similar to components and finished-goods sourcing, supplementary services were sourced globally, either by insourcing or outsourcing. The higher the asset specificity and the lower the transaction frequency of the supplementary services, the higher the global insourcing used. Finally, insourcing and offshoring of supplementary services were negatively related to the market performance of a service.

The fit perspective is therefore summarized as follows: There is a range of contingency factors (i.e., capital intensity, degree of service inseparability, market uncertainty, and transaction frequency) at the transaction-, firm-, and context-levels. These factors determine how much global outsourcing (both onshore and offshore) ought to take place from a performance perspective. To an extent, the contingency factors also explain how much global outsourcing actually takes place in practice. Fit is achieved when the actual global outsourcing level is in accordance with the level predicted

on the basis of the contingency factors. If a firm matches a global outsourcing decision to the relevant contingency factors, the resulting strategic fit helps achieve superior performance.

A "balanced-fit" perspective. The previous discussion raises two related questions. First, are these contradictory or rather complementary perspectives; if they are complementary, how do they complement each other? Second, how can we, taking into account these perspectives, explain the large increases in offshore and global outsourcing? We now seek to answer these two questions on the basis of the extant literature, specifically by drawing upon possible conceptual angles on global outsourcing.

To describe how the balance and fit perspectives complement each other, and to explain why over the past two decades or so we have witnessed the degree of global sourcing shifting toward more offshoring, we need to draw more directly upon key academic perspectives on global sourcing. We summarize 11 such perspectives in Table 4.

It is not in the scope of this article to describe each perspective in detail or to show how different perspectives are useful in predicting global outsourcing (for a more detailed description, see Mol, 2007). However, it is important to note that these perspectives operate at three different levels: the transaction, the firm, and the industry and institutional contexts. Taken together, they represent almost all the contingency factors that the academic literature has produced to date. Which of these perspectives matters most is to an extent determined by the empirical context in which outsourcing is investigated. Some of the perspectives have been more prominent than others in recent academic studies of outsourcing. Transaction-cost economics and the resource-based view come to mind as examples, which may reflect their actual importance in practice.

This takes us back to the two questions. The first question can be answered by stating that exactly where the optimal point of outsourcing (balance) lies is determined by the scores on the contingency factors (fit). In terms of the second question, the optimal point in terms of how much a firm should engage in offshoring will shift over time. Over the past two decades or so, we have witnessed that the degree of global sourcing has shifted to the right in Figure 1, that is, toward more outsourcing and offshoring. This implies that changes in both the level of the contingency factors as well as their constitution (i.e., which variables matter and to what extent) have caused the increase in outsourcing and offshoring levels.

Taken together, the implication is that the balance in global outsourcing has shifted toward higher levels of outsourcing because of the need to fit global outsourcing levels to a set of changed circumstances. We suggest two major drivers of this change. First, IT, including the Internet, has greatly facilitated cross-border business-to-business transactions. Second, institutional changes, which lie at the heart of the rise of both China and India as supply destinations, have also facilitated cross-border trade and investment, both of which in turn lead to more global outsourcing. These conclusions provide different implications for managers.

Conclusions and Managerial Implications

On the basis of these discussions, managers should rethink and redesign their global outsourcing activities. Many managers have a strong general sense for what constitutes a sound

Table 4 Perspectives on global outsourcing.

	Firm	*Context*	*Transaction*
Past	Resource-based view	Social networks	—
Present	Costly contracting Microeconomics Core competencies	Industrial organization Institutional voids	Transaction-cost economics Agency
Future	Real options Relations and learning	—	—

outsourcing and offshoring policy. They realize that outsourcing and offshoring every activity may lead to disasters, just as much as they recognize that not all activities should be insourced.

There is currently a tendency in practice to describe performance problems related to outsourcing or offshoring as *implementation issues*. Managers often assume that their outsourcing or offshoring decision is the proper design choice, and tend to attribute its unsatisfactory performance to various implementation problems that occur when dealing with independent and overseas suppliers. We suggest that there are many more fundamental problems of outsourcing or offshoring that are unrelated to implementation problems. Rather, there are limits to outsourcing and offshoring, and many inputs of a firm should not be outsourced or offshored.

Managers are often not conscious of the fact that there is an optimal degree of outsourcing across their entire portfolio (Leiblein, Reuer, and Dalsace, 2002). Instead of using this portfolio level, they tend to see the good or the evil of outsourcing or offshoring particular items or activities in that suppliers are not well equipped, insufficient guarantees are built into contracts, or market circumstances change rapidly. Many firms do not conduct enough analysis before they jump into outsourcing or offshoring. This helps explain why, in practice, outsourcing or offshoring often looks like a bandwagoning process. Likewise, many academic approaches have centered on analyzing individual make-or-buy decisions.

However, the performance advantages of outsourcing or offshoring will only materialize when a firm has the organizational capacity to integrate outsourced and/or offshored items/activities into its operations. Furthermore, many companies make outsourcing or offshoring decisions by evaluating only a few options on the basis of their previous experience and by what their competitors are doing.

Managers are in need of guidelines as to where the optimal point lies for their particular business at a particular time. On the basis of the contingency approach using a "fit" perspective, we can suggest several indicators to help answer that question including asset specificity, uncertainty, firm competencies, industry trends, and firm nationality and location. These moderating factors may help determine what is optimal for a particular firm at a particular time. Timing is crucial, as the optimal point will shift due to changes internal and external to the firm.

From a managerial perspective, developing a model that helps determine a firm's optimal degree of outsourcing or offshoring would be very useful. On the basis of this model, managers could prioritize their set of activities and outsource or offshore until they more or less reach optimality. As global sourcing is a dynamic process, competing firms may not accurately grasp the full benefit (and cost) of outsourcing or offshoring activities due to causal ambiguity. Simply bandwagoning on the first mover's current outsourcing or offshoring strategy offers no guarantee for improved performance. We suggest that tackling that challenge involves a broader behavioral understanding of how firms' outsourcing or offshoring trajectories change over time and within industries.

Acknowledgment

The authors acknowledge Michael J. Mol for his inputs on their previous research projects that help build the foundation of this article.

Bibliography

Bettis, R., Bradley, S., and Hamel, G. (1992) Outsourcing and industrial decline. *Academy of Management Executive*, **6** (1), 7–16.

Drazin, R. and Van de Ven, A.H. (1985) Alternative forms of fit in contingency theory. *Administrative Science Quarterly*, **30**, 514–539.

Dunning, J.H. (1993) *Multinational Enterprises and the Global Economy*, Addison-Wesley, Wokingham.

Gottfredson, M., Puryear, R., and Phillips, S. (2005) Strategic sourcing – from periphery to the core. *Harvard Business Review*, **83** (2), 132–139.

Kotabe, M. (1992) *Global Sourcing Strategy: R&D, Manufacturing, and Marketing Interfaces*, Quorum Books, New York.

Kotabe, M. (1998) Efficiency vs. effectiveness orientation of global sourcing strategy: a comparison of U.S. and Japanese multinational companies. *Academy of Management Executive*, **12** (4), 107–119.

Leiblein, M.J., Reuer, J.J., and Dalsace, F. (2002) Do make or buy decisions matter? The Influence of organizational governance on technological performance. *Strategic Management Journal*, **23** (9), 817–833.

Miles, R.E. and Snow, C.C. (1986) Organizations: new concepts for new firms. *California Management Review*, **28** (Spring), 62–73.

Mol, M.J. (2007) *Outsourcing: Design, Process, and Performance*, Cambridge University Press, Cambridge.

Murray, J.Y. and Kotabe, M. (1999) Sourcing strategies of U.S. service companies: a modified transaction-cost analysis. *Strategic Management Journal*, **20**, 791–809.

Murray, J.Y., Kotabe, M., and Wildt, A.R. (1995) Strategic and financial implications of global sourcing strategy: a contingency analysis. *Journal of International Business Studies*, **26** (1), 181–202.

Murray, J.Y., Kotabe, M., and Zhou, J.N. (2005) Strategic alliance-based sourcing and market performance: evidence from foreign firms operating in China. *Journal of International Business Studies*, **36** (2), 187–208.

Quinn, J.B. and Hilmer, F.G. (1994) Strategic outsourcing. *Sloan Management Review*, **35** (4), 43–55.

Rangan, S. (2000) The problem of search and deliberation in international exchange: Microfoundations to some macro patterns. *Journal of International Business Studies*, **31** (2), 205–222.

Trent, R.J. and Monczka, R.M. (2003) International purchasing and global sourcing – what are the differences? *Journal of Supply Chain Management*, **39** (4), 26–37.

Williamson, O.E. (1985) *The Economic Institutions of Capitalism*, Free Press, New York.

export performance

Luis Filipe Lages and Carlos M. P. Sousa

Introduction

The area of export performance is attracting both academic and managerial attention at an increasing pace. For more than three decades (1964–1998), work in the field was mostly concerned with understanding export-performance determinants. With the publication of a special issue on export performance measurement in the *Journal of International Marketing* in 1998, research in this topic received a new boost. Since then, several works have been presented to justify the existence of divergent findings in this field (Lages and Lages, 2004; Lages, Lages, and Lages, 2005; Diamantopoulos and Kakkos, 2007). Simultaneously, new works appeared arguing that export managers are not only proactive but also reactive to past results. As a consequence, research should also start considering export performance as an independent variable (Lages *et al*., 2008b).

Today, after almost half a century of research in this theme, there is very little consensus on the key antecedents and outcomes of export performance as well as how export performance should be defined and measured (Sousa, 2004; Sousa, Martinez-Lopez, and Coelho, 2008). Not surprisingly, the current literature on export performance is fragmented, diverse, and inconsistent, hindering advancement in the field. It is fragmented because of numerous studies in the literature that are characterized for adopting a variety of analytical techniques and methodological approaches. It can be classified as diverse because of the different determinants and measures of export performance (*see* STRATEGIC EXPORT MARKETING–ACHIEVING SUCCESS IN A HARSH ENVIRONMENT). Finally, it is considered inconsistent because of the different and often contradicting findings that exist in the export-performance literature. Despite the critical importance of this theme both from public policy making and managerial perspectives, in our view, no consensus will exist in the literature until an established measure of export performance is used. The purpose of this article is to provide an overview of the literature on export performance and a list of potential directions for future research.

Subjective or Objective Measures?

Over the last five decades, a wide diversity of measures has been used in the exporting literature to assess performance (Diamantopoulos and Kakkos, 2007). A first choice to be made is between subjective and objective measures (*see* MARKETING METRICS). Supporters of subjective assessment argue that, although objective assessments in measuring actual export performance may be regarded as being "more trustworthy," these measures raise different measurement problems. Obtaining accurate objective data on export performance is very hard because export managers are rarely willing to respond effectively to absolute values (Lages, Lages, and Lages, 2005). They will also argue that managers often disagree about which operational measures to use when setting targets because performance assessment is idiosyncratic to the type of firm and its environment, and some measures (e.g., profitability, sales, and ROI) raise comparability problems due to different accounting practices within strategic business units (SBUs) and across firms (Styles, 1998). Another major obstacle is that financial reports hardly ever make a distinction between the performance of domestic and export markets operations and even more seldom provide data on each export venture. As a consequence, objective metrics are rarely used because it is often impossible to establish a common definition of success/failure or fixed reference points across firms. On the other hand, by measuring perceived performance, researchers are able to capture the degree to which performance has matched the aspiration levels of the firm from one year to the next. The imaginary line separating success from failure will be determined by the level of expected performance and will serve as a useful starting point for future decision making. In sum, one may assume that an individual export venture will be successful when targets are met or exceeded and unsuccessful when it is below this *line* (Lages *et al*., 2008b). Organizational learning theory posits that

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

performance achievement and satisfaction are crucial in triggering action, because managers are assumed to have a set performance goals to which performance outcomes are compared. Subjective measures of performance capture the degree to which performance matched these goals and aspiration levels of the firm, and seem to be directly amenable to comparison across firms. Managerial perceptions of the achievement of these exporting goals and satisfaction with the exporting activity may play a key role in the definition of firms' export activities (Lages *et al.*, 2008b). Future research should consider both objective and subjective dimensions of export performance (Katsikeas, Leonidou, and Morgan, 2000; Sousa, 2004).

Aggregated or Desegregated Measures?

Past research suggests that no single measure is adequate to assess export performance. The use of multiple measures and ways to measure export performance is necessary to fully realize the strengths of each indicator and to minimize the impact of their shortcomings (Sousa, 2004). No approach is perfect, and through a combination of several approaches one may increase the likelihood of having solid findings. Hence, for testing the robustness of export-performance findings, future research is strongly encouraged to use a combination of different approaches: total item aggregation, intermediary aggregation, and total desegregation approaches. While in the most aggregated approach, findings result in a loss of specific information, because the distinction among the independent items is lost, in the most disaggregated approach, it becomes harder to have an overall picture of export performance. As a consequence, the most common approach to assess export performance is an intermediary aggregation of various items into single performance measures (i.e., commonly denominated as *factors* or *constructs*) (Lages, Lages, and Lages, 2005). Although this combination of different measures into factors allows better contextualizing the complex performance construct (Katsikeas, Leonidou, and Morgan, 2000), this approach does not precisely capture the different dimensions of the performance phenomenon (Doyle and Stern, 2006). In addition, some measures may correlate differently with other measures and different measures may have different weights depending on the context. For example, in the case of Western companies, rankings for specific performance variables have been created and profits became one of the most important performance measures. However, when managers focus on seeking outstanding profits, this may lead to poor results in other export performance metrics such as market share (Doyle and Stern, 2006). Likewise, when managers focus too much on sales volume, this might create problems on sales revenue, and, as a consequence, create image problems. As such, for future research, we strongly advocate the use of aggregation approaches in combination with single-performance items. This way, by using the most disaggregated approach, it will also be possible to treat each one of the items included in the model individually (Doyle and Stern, 2006).

Static or Dynamic Approaches?

Although management research has been gradually moving from a static analysis to a more dynamic approach, research involving export performance predominantly follows a static approach. The literature on export performance rarely looks to change in terms of export performance and rarely implements a longitudinal research. This makes it difficult to implement dynamic models and limits efficacious measurement of performance (Katsikeas, Leonidou, and Morgan, 2000). Future longitudinal research allows a better analysis of the relationship between export performance and its determinants. Future research is strongly encouraged to build on organizational learning theory and strategic dynamics to better understand how export performance change occurs in foreign markets. Notwithstanding the significant amount of research conducted in exporting, a review of the literature indicates several shortcomings in this field.

Short- or Long-Term Assessment?

Although the majority of the literature assumes that managers undertake a proactive and

long-term perspective of export performance, a number of studies have suggested that western firms are often reactive and short-term oriented (Madsen, 1998; Lages and Lages, 2004). Most studies do not specify the time period and assume that export performance is a long-term issue (Diamantopoulos and Kakkos, 2007). Among the rare exceptions that do so, while some prefer to use a more dynamic approach by trying to capture performance over a 12-month period (Morgan, Kaleka, and Katsikeas, 2004; Lages and Lages, 2004), others follow a more static approach by assessing export performance levels in a specific year (Lages, Lages, and Lages, 2005). To organize the literature in terms of export performance measurement, future research should make this aspect clear. Indeed, short-term analysis is a critical issue from a managerial perspective as many top managers wish to *see* short-term performance effects and do not give enough time to observe the effects of strategy in the long term. However, being overly focused on short-term goals may be risky for the long-term development of a firm's capabilities (Madsen, 1998). The importance of performance dimensions may also vary across stakeholder groups (e.g., investors, employees, and customers) and depend on whether the focus is on the short term or the long term (Sousa, 2004). A manager who focuses on the long term to increase the market share in a foreign market may not perceive export performance to be low when export sales or export profits are weak.

A Dependent or an Independent Variable?

Interestingly, and despite the fact that the overwhelming majority of studies analyze performance as a dependent variable, it is uncertain which measures should be analyzed as causally dependent. Researchers tend to put export performance as a dependent variable even when the data collected on strategy and performance variables relate to the same period of time. However, managerial practice is expected to be dynamic. Past performance is highly likely to influence organizational change as managers tend to respond to performance feedback. However, to our knowledge, no study has analyzed the impact of past performance on strategic change in international markets. Performance as an independent variable has been highly unexplored in an international marketing context (Lages and Montgomery, 2004; Lages *et al.*, 2008b are two exceptions). As the underlying aspects of export-marketing strategy are driven by managerial action, greater understanding of managerial learning from past export performance can provide marketing academics and practitioners with strategic insights into enhancing export performance (*see* CONCEPT OF CAUSALITY AND CONDITIONS FOR CAUSALITY).

Theoretical or Managerial Approaches?

In the late 1990s, some attempts were made to develop comprehensive and psychometrically sound measures of export performance (Styles, 1998; Zou, Taylor, and Osland, 1998). Despite these attempts, a major concern remained that export-performance assessment tends to be misaligned with the managerial world (Madsen, 1998; Lages and Lages, 2004). Several reasons justify the need for a sound managerial evaluation of export performance. If one considers that exporting actions have an impact on the overall failures/successes of firms, it becomes necessary to have tools that allow managers to monitor export performance. Proper measures will provide decision makers with a reference to support the definition of future short- and long-term actions, such as the allocation of human and financial resources to specific export ventures. Export performance evaluation is equally important at the public-policy level. The benefits provided by the exporting activity encourage public policy makers to implement export assistance programs to enhance firms' global marketing strategies. As a consequence, a proper assessment of export-marketing strategy (Lages, Abrantes, and Lages, 2008a) and export performance will have an impact on any country's economic health (Lages and Montgomery, 2005).

During the last four years, some export measurement tools have been developed with greater managerial- and public-policy flavor. One of these tools is the STEP scale (Lages and

Lages, 2004), which addresses the short-time horizon frequently used by managers and public-policy makers to assess performance. Two other managerial tools – the APEV scale and the PERFEX scorecard (Lages, Lages, and Lages, 2005) – were specifically designed to be included in annual reports and to assess export performance both at the corporate and exporting venture levels. The APEV scale was inspired by the EXPERF measure (Zou, Taylor, and Osland, 1998). All these three scales (EXPERF, STEP, and APEV) are of a reflective nature and were developed using CONFIRMATORY FACTOR ANALYSIS. More recently, Diamantopoulos and Kakkos (2007) suggest that this field should also use formative indicators. As such, the authors propose the AEP index as a composite measure of managerial-subjective evaluations of export performance. This index enables the assessment and comparison of multiple export objectives, both within and between firms, by allowing the setting of the weights of the different indicators to be zero. From a managerial perspective, the major advantage of using performance metrics of a formative nature is that, in contrast to reflective aggregated measures, it becomes possible to identify which particular indicators the managers have in mind when assessing export performance.

Determinants of Export Performance

Two broad theoretical approaches, the resource-based paradigm and the contingency paradigm, provide the basis for classifying the determinants of export performance into internal and external factors. Specifically, internal determinants are justified by resource-based theory, while external determinants are supported by contingency theory. Resource-based theory focuses on how sustained competitive advantage is generated by the unique bundle of resources at the core of the firm. The contingency paradigm suggests that environmental factors influence the firm's strategies and export performance (*see* STANDARDIZATION/ADAPTATION OF INTERNATIONAL MARKETING STRATEGY). The effects of various firm characteristics on export performance are dependent on the specific context of the firm. An extensive list of studies has already identified key determinants of success in terms of the internal and external factors influencing successful export performance (Sousa, Martinez-Lopez, and Coelho, 2008). However, there is a lack of agreement on the domains and measurement of the determinants of export performance. This obstructs development of the theory on export performance, and makes it very difficult to compare the findings from different studies and literature. As a result, attempts should be made to develop clear conceptual domains and sound schemes to measure the variables. Nonetheless, after more than four decades of research on the analysis of the relationships among internal and external forces, export-marketing strategy and export performance, researchers now agree that export performance must be analyzed as a function of the fit between the firm's environment and the selected export-marketing strategy.

Despite the argument that control variables deserve as much attention and respect as do independent and dependent variables, most export performance studies fail to include them (Sousa, Martinez-Lopez, and Coelho, 2008). This disregard for the role of control variables is an issue of concern, and researchers are encouraged to account for these effects in future studies. In addition to the analysis of possible direct relationships (as undertaken by most investigators), future research is also encouraged to analyze the moderating effects of external forces and the indirect impact of environmental forces on export performance through their influence on strategy. The export-performance literature has reached a level of sophistication that researchers should be interested in detecting not only the main effects of independent variables but also their indirect and moderating effects.

Unit of Analysis

There is no consensus in the literature regarding the level of performance assessment. Most export studies have looked at export performance at the firm level (Sousa, Martinez-Lopez, and Coelho, 2008). One possible explanation for this predilection by researchers could be the fact that respondents are more willing to disclose information at this broad level. The underlying

theoretical justification for firm-level studies is the theory on internalization (Rugman, 1980). This theory states that, in imperfect markets, firms should internalize the firm-specific advantages, both tangible and intangible, to extract maximum economic rent. Consequently, the study of export performance at the firm level has the benefit of capturing firm-specific advantages, which are derived not only from the development of a particular product but also from the total learning process of the firm. This allows examine the influence of potential determinants (e.g., overall firm strategy, organizational culture, organizational structure, R&D, etc.) that are not directly related to a specific venture. Other units that are commonly used are at the product level, strategic business-unit level, and product–market export venture level. Proponents of these three levels argue that it is unrealistic to expect that the same strategies can lead to the same results in all export products, SBUs, and product–market ventures. Naturally, the use of different levels of analysis will lead to different (and sometimes conflicting) insights on the topic.

Frames of Reference and Directions for Future Research

Researchers use different frames of reference. In most cases, performance measurement tends to be self-driven. More recently, some works have explored performance assessment versus competitors (Morgan, Kaleka, and Katsikeas, 2004). Future research should consider the use of customer-, stakeholder-, and network-driven perspectives. In addition, employees within the same firm will hold different viewpoints of the same reality. As a consequence, the use of multiple informants within each firm will bring added value to assess this complex phenomenon (Sousa, 2004). With rare exceptions (Styles, 1998; Zou, Taylor, and Osland, 1998; Lages and Lages, 2004), most studies have used a single country or region as a frame of reference. The performance measures used in these studies often reflect the unique emphasis that different countries place on exporting (Zou, Taylor, and Osland, 1998). As a result, attempts should be made to validate scales across countries. However, considerable difficulties are likely to be encountered in establishing equivalence and comparability of research in different studies. Researchers have to develop cross-cultural conceptualization and measurement that reflect true cultural differences among markets along the underlying construct under study. This can play an important part in advancing export-marketing theory by stimulating cross-cultural export-marketing studies that investigate specific similarities and differences among and between countries (Styles, 1998).

The United States is the most researched country in export-performance studies and despite a rise in the number of studies conducted outside the United States, there are still countries from certain parts of Asia, South and Central America, the Caribbean, and Africa that have received little or no attention from researchers (Sousa, Martinez-Lopez, and Coelho, 2008). Further research should consider the inclusion of such countries to investigate whether our current knowledge can be generalized to these countries, especially those from the developing world. Firms from developing countries are particularly interesting to study in future research because of their growing presence in an integrated global economy. In addition, most export performance studies involve samples drawn form manufacturing industries with relatively few studies investigating export performance of service firms. While there are some determinants and measures of export performance that apply to both manufacturing goods and services, it is likely that additional variables must be taken into account that relate to the specific characteristics of services firms when operating in the international arena (Sousa, Martinez-Lopez, and Coelho, 2008). Considering that services account today for around 20–30% of world trade, there is an increasing need for researchers to test whether traditional theories of GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES apply to the international marketing of services.

Taken together, the existing shortcomings in the export-performance literature create new research opportunities as they leave international marketing academics and practitioners without a clear understanding of the effects of

export performance as well as of the factors influencing it.

ACKNOWLEDGMENT

This work was funded by "Fundação para a Ciência e a Tecnologia" and NOVA FORUM.

Bibliography

Diamantopoulos, A. and Kakkos, N. (2007) Managerial assessments of export performance: conceptual framework and empirical illustration. *Journal of International Marketing*, **15** (3), 1–31.

Doyle, P. and Stern, P. (2006) *Marketing Management and Strategy*, 4th edn, FT Prentice Hall.

Katsikeas, C.S., Leonidou, L.C., and Morgan, N.A. (2000) Firm-level export performance assessment: review, evaluation, and development. *Journal of the Academy of Marketing Science*, **28** (4), 493–511.

Lages, L.F., Abrantes, J.L., and Lages, C.R. (2008a) The STRATADAPT scale: a measure of marketing strategy adaptation to international business markets. *International Marketing Review*, **25** (5), 584–600 (Special Issue Business-to-Business as International Business).

Lages, L., Jap, S., and Griffith, D. (2008b) The role of past performance in export ventures: a short-term reactive approach. *Journal of International Business Studies*, **39** (2), 304–325.

Lages, L.F. and Lages, C.R. (2004) The STEP scale: a measure of short-term export performance improvement. *Journal of International Marketing*, **12** (1), 36–56.

Lages, L., Lages, F.C., and Lages, C.R. (2005) Bringing export performance metrics into annual reports: the APEV scale and the PERFEX scorecard. *Journal of International Marketing*, **13** (3), 79–104.

Lages, L.F. and Montgomery, D.B. (2004) Export performance as an antecedent of export commitment and marketing strategy adaptation: evidence from small and medium sized exporters. *European Journal of Marketing*, **38** (9/10), 1186–1214.

Lages, L.F. and Montgomery, D.B. (2005) The relationship between export assistance and performance improvement in Portuguese export ventures: an empirical testing of the mediating role of pricing strategy adaptation. *European Journal of Marketing*, **39** (7/8), 755–784.

Madsen, T.K. (1998) Executive insights: managerial judgment of export performance. *Journal of International Marketing*, **6** (3), 82–93.

Morgan, N.A., Kaleka, A., and Katsikeas, C.S. (2004) Antecedents of export venture performance: a theoretical model and empirical assessment. *Journal of Marketing*, **68** (1), 90–108.

Rugman, A.M. (1980) A new theory of the multinational enterprise: internationalization versus internalization. *Columbia Journal of World Business*, **15** (1), 23–29.

Sousa, C.M. (2004) Export performance measurement: an evaluation of the empirical research in the literature. *Academy of Marketing Science Review* **8** (9), 1–22. Available: http://www.amsreview.org/articles/sousa09-2004.pdf.

Sousa, C.M.P., Martinez-Lopez, F.J., and Coelho, F. (2008) The determinants of export performance: a review of the research in the literature between 1998 and 2005. *International Journal of Management Reviews*, **10** (4), 343–374.

Styles, C. (1998) Export performance measures in Australia and the United Kingdom. *Journal of International Marketing*, **6** (3), 12–36.

Zou, S., Taylor, C.R., and Osland, G.E. (1998) The EXPERF scale: a cross-national export performance measure. *Journal of International Marketing*, **6** (3), 37–58.

forces affecting global integration and global marketing

Kate Gillespie

Global integration is the state in which one national market (including a firm's home market) is significantly affected by other national markets. Because of global integration, international marketers cannot continue to successfully employ multiple single-market strategies but instead must adopt GLOBAL MARKETING STRATEGY that address multiple, if not all, national markets at the same time. Examples of such global strategies include global rollouts of new products, centralized production of standardized products, and global-theme advertising. In the last quarter of the twentieth century, several forces, some interrelated, resulted in greater global integration. These forces have been categorized in different ways but generally fall into five categories: technological, buyer behavior, competitive, cost, and regulatory. Each industry is affected differently by the forces of global integration, but virtually all industries are affected in some way by these forces.

Levitt (1983) first identified the emergence of global integration, focusing on the role of technology and the convergence of buyer behavior in what he termed the *globalization of markets*. Levitt argued that technological advancements, especially in the areas of communication, transport, and travel, made buyers in disparate lands aware of the same modern products. As a result, consumer wants and needs across countries had become irrevocably homogenized (*see* GLOBAL CONSUMERISM AND CONSUMPTION). This allows global marketers to sell more standardized products across different national markets (*see* STANDARDIZATION/ADAPTATION OF INTERNATIONAL MARKETING STRATEGY). Later authors have elaborated on both the role of technology and buyer behavior in global integration. For example, Yip (2003) notes that the speed at which new technologies become obsolete suggests that new products should be introduced quickly into all major markets at once to capture global-scale profits to offset shortened product life cycles (*see* STAGES OF THE PRODUCT LIFE CYCLE). He also argues that the increase in per capita income in developing countries allows consumers in such countries to aspire to and afford products similar to those purchased in developed countries.

Competitive forces also drive global integration and, consequently, global marketing strategies. A new awareness of and need for foreign markets on the part of firms, combined with the convergence of buyer behavior across markets, has intensified global competition. On the one hand, many firms can more easily enter foreign markets without extensive product adaptation. On the other hand, foreign competitors can more easily enter the domestic market of those same firms. Thus global marketing has come to encompass strategic intelligence of global competitors. Firms increasingly study and respond to global competitors while acknowledging that their competitors are doing the same (*see* COMPETITOR ANALYSIS). Thus another argument for the global rollout of products is the need to preempt global competitors from stealing new product ideas introduced in a firm's home market before they are introduced in other markets. In addition, global competitive moves, such as attacking a competitor in their home market as retaliation for an attack on one's own domestic market, have become increasingly important to global marketers (Yip, 2003).

As global competition intensifies, price competition becomes increasingly important in most industries (*see* INTERNATIONAL PRICING OBJECTIVES AND STRATEGIES). As a result, cost considerations drive global integration as well. Standardizing products, or even partially standardizing products, across national markets allows firms to take advantage of production economies of scale. Centralized purchasing for needed inputs across national operations decreases operating costs in multinational firms. Global-theme advertising (*see* INTERNATIONAL ADVERTISING – IS THERE STILL A STANDARDIZATION VERSUS LOCAL ADAPTATION DEBATE?) cuts promotion costs worldwide by standardizing portions of a multinational campaign while allowing certain national adaptations (Gillespie, Jeannet, and Hennessey, 2007).

Perhaps the most enabling driver of global integration has been the widespread market

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

liberalization that has encompassed both developed and developing countries in the last quarter of the twentieth century. Bilateral and multilateral trade agreements allow global marketers to capture cost savings from centralized production by allowing access to foreign markets via exports subject to no or low tariffs. Similarly national regulations discouraging foreign acquisition of local firms have increasingly disappeared, allowing foreign competitors to quickly buy market share in national markets. Transnational standards have emerged relating to many aspects of international marketing such as product specifications, rules relating to bribery, and legislation protecting brands (Gillespie, Krishna, and Jarvis, 2002). Without such regulatory liberalization and standardization, the other drivers of global integration would have far less impact on global marketing.

Bibliography

Gillespie, K., Jeannet, J.P., and Hennessey, H.D. (2007) *Global Marketing*, Houghton Mifflin, New York.

Gillespie, K., Krishna, K., and Jarvis, S. (2002) Protecting global brands: toward a global norm. *Journal of International Marketing*, **10**, 99–112.

Levitt, T. (1983) The globalization of markets. *Harvard Business Review*, **61** 92–102.

Yip, G. (2003) *Total Global Strategy*, Prentice Hall, Upper Saddle River.

stages of market development

Peter Magnusson and Stanford A. Westjohn

A market's stage of development influences decisions relating to a firm's international marketing strategy, for example, entry mode, exporting, and GLOBAL SOURCING STRATEGY: AN EVOLUTION. However, categorizing a market depends on which classification is used. The classifications made by several notable international organizations are summarized in Table 1. The World Bank provides the strictest classification criteria. Countries are split into four categories based on gross national income (GNI) per capita, with the added provision that low- and middle-income economies are sometimes viewed as "developing," while high-income economies are considered "developed" (World Bank, 2008). The IMF identifies just two categories, "advanced," and "emerging and developing economies" (International Monetary Fund, 2008). The UN's UNCTAD arm uses the labels "developed," "economies in transition" (includes former Eastern European and Soviet-controlled countries), "developing," and "least developed countries" (UNCTAD, 2008). Although IMF's and UNCTAD's categories differ, both organizations maintain that strict criteria do not exist to determine a country's classification and that designations are for statistical convenience and do not necessarily express a judgment about a country's stage of development.

Stages of development need not be measured in purely economic terms, for example, GNI/GDP per capita, but may also include social and political development. For example, the United Nations Development Programme developed the Human Development Index (HDI). In addition to standard of living, captured by GDP per capita, the HDI includes life expectancy and access to knowledge, measured by literacy rates and education enrollment (UN Development Programme, 2008). The HDI reveals that some countries lag behind in terms of life expectancy and education despite having relatively high incomes. For example, the HDI rankings of South Africa and UAE are well behind their GDP/capita rankings. The Economist Intelligence Unit (EIU) has also developed indices that measure market development. The "quality-of-life" index uses material well-being (GDP/capita), health, political stability and security, family life, community life, climate and geography, job security, political freedom, and gender equality (*see* CONSUMER WELL-BEING; FORCES AFFECTING GLOBAL INTEGRATION AND GLOBAL MARKETING) (The Economist, 2005). A second EIU measure, the Index of Democracy, focuses exclusively on a country's political progress toward democracy along five categories: electoral process and pluralism, civil liberties, functioning of government, political participation, and political culture (The Economist, 2008). The Index of Democracy reports a positive correlation between democracy and GDP/capita, and research suggests that the primary direction of causation is from democracy to income (Rigobon and Rodrik, 2005).

There are several methods for measuring market development but no universal standard exists. Traditionally, market development is based on economic development (i.e., GDP/capita), but more recent measures attempt to incorporate additional dimensions (i.e., social and political development). Notably, the popular "EMERGING MARKETS" label is largely absent from official categorizations, with the exception of IMF, which labels the "developing" category as "emerging and developing" markets. Furthermore, while economic growth indicates expanding commercial opportunities (*see* GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES; GLOBAL MARKETING STRATEGY), EIU found that economic growth is not related to a population's level of life satisfaction (The Economist, 2005). Given the inconsistent labeling of stages of development, researchers must carefully assess what is being measured when comparing more than one report or study. In making stand-alone evaluations of market development, researchers should select a classification scheme that best suits the question being asked.

Bibliography

International Monetary Fund (2008) World Economic Outlook 2008 (Online). Available: http://www.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 Stages of market development.

World Bank	*International Monetary Fund*	*UN (UNCTAD)*	*EIU's Index of Democracy*	*UN (UNDP) Human Development Index*	*EIU's Quality-of-life Index*
High income (>$11 456)	Advanced economies (including newly industrialized Asian economies)	Developed	Full democracies	High human development (>0.8)	High quality of life (>7)
Upper-middle income ($3706–11 455)	Other emerging markets and developing countries	In transition	Flawed democracies	Medium human development (0.5–0.79)	
Lower-middle income ($936–3705)		Developing	Hybrid regimes		
Low income (<$935)		Least developed	Authoritarian regimes	Low human development (<0.50)	

imf.org/external/pubs/ft/weo/2008/02/index.htm (28 February 2009).

Rigobon, R. and Rodrik, D. (2005) Rule of law, democracy, openness, and income: estimating the interrelationships. *Economics of Transition*, **13**, 533–564.

The Economist (2005) The Economist Intelligence Unit's Quality-of-Life Index (Online). Available: http://www.economist.com/media/pdf/QUALITY_OF_LIFE.PDF (28 February 2009).

The Economist (2008) The Economist Intelligence Unit's Index of Democracy 2008 (Online). Available: http://a330.g.akamai.net/7/330/25828/20081021185552/ graphics. eiu. com/ PDF/ Democracy%20 Index% (3 March 2009).

UNCTAD (2008) Handbook of Statistics 2008 (Online). Available: http://www.unctad.org/statistics/handbook (28 February 2009).

UN Development Programme (2008) Human Development Indices: A Statistical Update 2008 (Online). Available: http://hdr.undp.org/en/mediacentre/news/title,15493,en.html (28 February 2009).

World Bank (2008) World Development Indicators 2008 (Online). Available: http://go.worldbank.org/U0FSM7AQ40 (28 February 2009).

services globalization

Haksin Chan

With diminishing trade barriers (*see* EMBARGOES AND SANCTIONS), emerging technologies (*see* DIGITAL MEDIUM AND GLOBAL MARKETING), and converging tastes (*see* GLOBAL CONSUMERISM AND CONSUMPTION), the era of globalization has arrived. Even "unlikely global candidates" in highly regulated service sectors have joined the ranks of global competitors (Johansson, 2002). However, global expansion (*see* MARKET ENTRY AND EXPANSION) poses major challenges for service operations, which are sensitive to sociocultural factors because of the prominent human elements. The adage "think globally, act locally" is especially applicable to services marketers.

The Inseparability Problem

Some services require a higher degree of local adaptation than others (*see* STANDARDIZATION/ADAPTATION OF INTERNATIONAL MARKETING STRATEGY). Lovelock and Yip's (1996) categorization of services as people-processing (e.g., lodging), possession-processing (e.g., delivery), and information-based (e.g., insurance) highlights their different needs for adaptation. Typically, people-processing services (which involve customer participation throughout the service process) have the greatest need for localized operations. In general, services that are separable (i.e., services that can be performed away from the customer) enjoy more flexibility in their global expansion strategies than those that are inseparable (Erramilli and Rao, 1993). Since many services are characterized by inseparability (which precludes the less costly modes of entry), global expansion is usually a riskier choice for services marketers than for marketers of tangible goods. This may be an important reason for the *relatively* slow growth of global services (Grönroos, 1999). Nonetheless, innovative electronic- and self-service technologies should help circumvent the inseparability problem to some extent (*see* SERVICE INNOVATION MANAGEMENT).

The Variability Issue

Because of the inherent variability of service performance, service failures are common occurrences (*see* SERVICES MARKETING STRATEGY). In general, consumers in Eastern, collectivist societies are more tolerant of service failures than their Western, individualist counterparts – this is due to longstanding cultural differences and contemporary economic conditions (Zhang, Beatty, and Walsh, 2008). However, the pattern is reversed when the nature of failure is "social" rather than "economic" (Chan and Wan, 2008). This reversal is consistent with the anecdotal evidence that Asian and Western airlines emphasize social (e.g., attentive flight attendants) and economic (e.g., convenient schedules) attributes, respectively. A similar pattern regarding social and economic attributes also manifests in cross-cultural consumer responses to post-failure recovery efforts (Hui and Au, 2001). Apparently, culture is an important consideration for the control of service quality and the design of service-recovery programs in global markets (*see* CROSS-CULTURAL PSYCHOLOGY OF CONSUMER BEHAVIOR).

The Protectionism Concern

Although GATT/WTO negotiations have opened doors around the world, the pace of liberalization is slower for services than for manufactured goods. In particular, nontariff barriers for services are still pervasive (Javalgi and White, 2002). These barriers are formidable in sectors that involve strong vested interests at the national (e.g., banking) or professional (e.g., medical practice) level. In times of economic downturn, a resurgence of protectionism is a distinct possibility (Clark and Rajaratnam, 1999). On the consumer side, ethnocentrism (Shimp and Sharma, 1987), patriotism (Han, 1988), and animosity toward other nations (Klein, Ettenson, and Morris, 1998; *see* CONSUMER ANIMOSITY) are powerful forces that may hamper globalization. Note that the high visibility of many global service operations makes them easy targets in social and political crises. In 2008, Carrefour stores were the primary targets for boycott and protest when nationalistic Chinese expressed their outrage

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

at the French government. Arguably, the much-publicized fear of "McDonaldization" reflects a negative attitude toward the fast-food model as well as the Americana it represents.

The Need for Research

Practice has far outpaced research in global services, resulting in a significant knowledge gap. Until recently, the services marketing literature has devoted little attention to international issues, and the international marketing literature has seldom taken a service-oriented perspective. On a positive note, global services research is at the "takeoff stage" of its life cycle (Furrer and Sollberger, 2007). At the macro level, the continued trend of "servicization" should prompt refinement of existing global marketing theories (*see* GLOBAL MARKETING STRATEGY; GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES). At the micro level, research opportunities abound – for pinpointing culture-specific service quality dimensions, (dis)satisfaction responses, (*see* CUSTOMER SATISFACTION/DISSATISFACTION; CUSTOMER SATISFACTION) and recovery strategies.

Bibliography

Chan, H. and Wan, L.C. (2008) Consumer responses to service failures: a resource preference model of cultural influences. *Journal of International Marketing*, **16**, 72–97.

Clark, T. and Rajaratnam, D. (1999) International services: perspectives at century's end. *Journal of Services Marketing*, **13**, 298–310.

Erramilli, M.K. and Rao, C.P. (1993) Service firms' international entry-mode choice: a modified transaction-cost analysis approach. *Journal of Marketing*, **57**, 19–38.

Furrer, O. and Sollberger, P. (2007) The dynamics and evolution of the service marketing literature: 1993–2003. *Service Business*, **1**, 93–117.

Grönroos, C. (1999) Internationalization strategies for services. *Journal of Services Marketing*, **13**, 290–297.

Han, C.M. (1988) The role of consumer patriotism in the choice of domestic versus foreign products. *Journal of Advertising Research*, **28**, 25–32.

Hui, M.K. and Au, K. (2001) Justice perceptions of complaint-handling: a cross-cultural comparison between PRC and Canadian customers. *Journal of Business Research*, **52**, 161–173.

Javalgi, R.G. and White, D.S. (2002) Strategic challenges for the marketing of services internationally. *International Marketing Review*, **19**, 563–581.

Johansson J.K. (2002) Global marketing: research on foreign entry, local marketing, global management, in *Handbook of Marketing* (eds B. Weitz and R. Wensley), Sage Publications, London, pp. 457–483.

Klein, J.G., Ettenson, R., and Morris, M.D. (1998) The animosity model of foreign purchase: an empirical test in the People's Republic of China. *Journal of Marketing*, **62**, 89–100.

Lovelock, C.H. and Yip, G.S. (1996) Developing global strategies for service business. *California Management Review*, **38**, 64–86.

Shimp, T.A. and Sharma, S. (1987) Consumer ethnocentrism: construction and validation of the CETSCALE. *Journal of Marketing Research*, **24**, 280–289.

Zhang, J., Beatty, S.E., and Walsh, G. (2008) Review and future directions of cross-cultural consumer services research. *Journal of Business Research*, **61**, 211–224.

gray markets

Kersi D. Antia

Gray marketing – the unauthorized distribution of genuine branded products – is a worldwide occurrence. The gray market for technology products alone is estimated to be worth US$58 billion (KPMG-AGMA Gray Market Study, 2009). This is merely the tip of the iceberg, however, as gray marketed products span a wide spectrum, ranging from watches, pharmaceutical products, and running shoes, to commodity chemicals, automobiles, agricultural harvesters, and heavy earth moving equipment. It is widely acknowledged that any manufacturer with multinational ambitions or presence is more than likely to have faced gray marketing in some form or another (*see* MARKETING CHANNEL STRATEGY).

Accounts of gray marketing in the academic and trade press alike describe it as a single, monolithic phenomenon. The typical storyline is as follows – somewhere along the value chain (*see* MULTICHANNEL MARKETING), genuine product is "diverted" by authorized channel members to those outside the manufacturer-approved dealer network and ultimately to unwitting end-customers, who only realize the product is gray marketed when they need warranty support. The truth, however, is a little more subtle than that black and white characterization. In fact, gray markets occur under at least two strikingly contrasting circumstances, and are frequently participated in by knowledgeable, deal-prone businesses and end-customers that are well aware of the "gray market" status of the goods sold. Consider the two situations below:

In the first few days following the launch of the Apple iPhone in China (*see* LAUNCH STRATEGIES; MARKET ENTRY AND EXPANSION), less than 5000 units were sold (Hughes, 2008), a stark contrast to the heady reception the iPhone was accorded in its launch across several countries in Europe. The Chinese debacle was not altogether surprising, however, given that somewhere between 800 000 and 1.4 million iPhones are estimated to have been "unlocked" or able to run on networks other than Apple's designated partners (Morrison, 2008) by way of gray market channels, "... a by-product of *pent-up demand* for a much-hyped device (*see* BRAND EQUITY; TRADEMARKS, PROPRIETARY MARKS, AND BRANDS) made by a company *that places strict limits* on where and how it is sold" (Burrows, 2008). Moreover, the gray marketed iPhones in China had additional features (wi-fi capabilities, stripped from the officially sold product to comply with Chinese government regulations) at a significantly cheaper price (*see* PRICING STRATEGY)! Not surprisingly, the diversion of Apple's iconic product poses significant challenges for the company – US$300–500 million in foregone revenues and profits for every 1 million iPhones diverted for "unlocking" (Elmer-De Witt, 2008).

Whereas the preceding instance of gray marketing is attributable to the strong cachet of the iPhone and the perception of its limited availability, (Inman *et al.*, 1997) that is, a high demand-lagging supply condition, gray markets also arise when demand–supply conditions are the obverse. Thanks to an economy in recession, nearly 34 000 businesses filed for bankruptcy in 2008, a 42% increase from just a year earlier (Hoffman, 2008). With resellers unable to sell products in these challenging economic conditions despite markdowns of up to 70%, a glut of liquidation-fueled gray market product was only to be expected (Crosby, 2009). Manufacturers found themselves in the uncomfortable position of having to compete with their own products being sold at steep discounts!

It should be clear from the preceding examples that gray market activity may be fueled by strong demand and scarce supply, as well as trenchant sales and a supply glut. Moreover, across both cases, customers are not the "surprised and confused victims" as often portrayed; rather, they are knowledgeable, deal-prone, and ever-resourceful parties that demand the best value for their spending dollar. Firms cognizant of these differences must apply the corrective strategies appropriate to each situation, rather than adopt a "one size fits all" approach to managing gray markets.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Bibliography

Burrows, P. (2008) Inside the iPhone gray market. *BusinessWeek*, February 12. http://www.businessweek.com/ technology/ content/ feb2008/ tc20080211_152894.htm

Crosby, J. (2009) Retail Woes bring Cheer to Liquidators, Star Tribune, January 4. http://www.startribune.com/ business/ 37018479.html?elr=KArksLckD8EQDUoaEyqyP4O: DW3ckUiD3aPc: _Yyc: aUUsT (accessed 17 January 2010).

Elmer-De Witt, P. (2008) Apple's $300 million Gray Market Dilemma. http://brainstormtech.blogs.fortune.cnn. com/ 2008/01/28/apples- 300-million - gray- market- dilemma/ (accessed 15 December 2009).

Hoffman, K. (2008) Gray Market Swells as Economy Swoons, The Seattle Times, December 8. http:// seattletimes. nwsource. com/ html/ businesstechnology/2008480576_bttechliquidator08.html (accessed 17 January 2010).

Hughes, N. (2008) Despite Disappointing China Debut, iPhone's 2010 Predicted to be Strong. http://www.appleinsider. com/ articles/ 09/11/03/ despite_ disappointing_ china_ debut_ iphones_ 2010_predicted_to_ be_ strong.html (accessed 17 January 2010).

Inman, J.J., Anil, C.P. and Raghubir, P. (1997) Framing the deal: the role of restrictions in accentuating deal value. *Journal of Consumer Research*, **24** (1), 68–79.

KPMG-AGMA Gray Market Study (2009) KPMG-AGMA Survey of Gray Marketing of Technology Products http://us.kpmg.com/RutUS_prod/Documents/8/GrayMarket.pdf (accessed 18 December 2009).

Morrison, D. (2008) iPhone Shortage: Runaway Gray Market in Emerging Markets to Blame?. http:// moconews. net/ article/ 419- iphone- shortage- runaway- gray- market- in- emerging- markets- to- blame/ (accessed 17 December, 2009).

pricing and currency fluctuations in international marketing

Terry Clark, Daniel Rajaratnam, and David Campbell

As a topic of practical and academic concern, pricing is jointly claimed by the disciplines of finance as well as marketing. Both disciplines share an interest in how price relates to profit, and to competitive interaction. However, the interests of marketing go beyond this, extending to price as a component of value (*see* PRICING STRATEGY), as a communicator of quality, as a psychological cue (*see* HOW CONSUMERS RESPOND TO PRICE INFORMATION), and as a factor affecting brand equity (*see* PERCEPTION OF BRAND EQUITY).

In the international arena, a marketer's pricing concerns are complicated by two additional factors: (i) the existence of more than one currency and (ii) the existence of more than one culture (*see* BASE OF THE PYRAMID MARKETS: CULTURE INSIGHTS AND MARKETING IMPLICATIONS). We briefly explore the first of these.

The impact of multiple currencies on pricing decisions is clear when we consider the basic revenue and profit equations 1 and 2:

$$\text{Revenue} = \text{Price} \times \text{Quantity} \tag{1}$$

$$\text{Profit} = \text{Revenue} - \text{Costs} \tag{2}$$

In a pure export situation, in which costs are entirely in the currency of country A, and revenues entirely in that of country B, the profit equation becomes

$$P_{\mathrm{A}} = (R_{\mathrm{B}} \times FX_{\mathrm{AB}}) - C_{\mathrm{A}} \tag{3}$$

where

P_{A} = profit in country A's currency,
R_{B} = revenue in country B's currency,
C_{A} = costs in country A's currency, and
FX_{AB} = exchange rate between currencies A and B.

Here, in Equation 3 the exporter's profit margin straddles the exchange rate, and fluctuates with it, regardless of how efficiently the firm is run, or what marketing strategy (*see* GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES) is used.

Trade theory suggests that export (import) prices rise in response to trade imbalances. For example, suppose a world of just two trading nations A and B. Further suppose that nation A has a trade deficit with nation B. This implies strong selling pressure on nation A's currency, and a strong buying pressure on nation B's currency (i.e., importers in nation A must sell their own currency to buy nation B's currency to purchase nation B's goods). As this occurs, the price of nation A's currency goes down, and the price of nation B's increases (i.e., the exchange rate changes), resulting in a rise in the price of nation B's goods (from nation A's perspective), and a decrease in the price of nation A's goods (from nation B's perspective). The outcome is that nation A stops importing from nation B, and instead, begins to export to nation B, thus eliminating the trade deficit. Figure 1 illustrates this.

Here, the exporter from nation A sets price at time 1 in nation B at P_{B1}. This price determines the exporter's margin in nation A (M_{A1}). However, the exchange rate changes after time 1 (δFX_{AB1}). Trade theory (as explained above) suggests the exporter will adjust price in the export market each time the exchange rate changes, such that $M_{\mathrm{A1}} = M_{\mathrm{A2}} = M_{\mathrm{A}n}$. If this condition holds, trade imbalances will indeed be eliminated. However, trade theorists have noted the paradox that trade deficits persist, despite the solid logic of the argument outlined above. Why does this happen? Recent studies suggest that strategic actions of export marketing managers may, at least in part, explain the paradox.

What, for example, might we expect the firm to do when a currency appreciation/depreciation increases/decreases profit margins? A variety

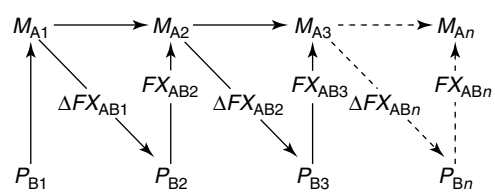


Figure 1 Pricing and profit margins in the presence of exchange rate changes.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

of options are possible. For example, with a favorable exchange rate change, the firm could either maintain host-country prices and pocket increased margins, or else lower host country price and increase market share, without losing revenue. Then again, if the firm wished to lower price so as to increase market share, they may be unable to do so because their existing distribution system cannot increase the flow of goods to the export market.

On the other hand, the firm experiencing an unfavorable exchange rate change (i.e., their margins are diminished), might want to raise host-country prices to regain margins. However, to do so could put them at a competitive disadvantage, and concede market share to competitors. More likely, they would maintain price, so as not to lose market share, and thus be stuck with diminished margins. Then again, if the firm had previously invested in brand equity, they might be able to increase price (and regain margins), without substantially affecting market share.

The situation becomes more complex yet when firms from different countries compete against each other in third-country export markets. In such situations, one firm could be disadvantaged by currency fluctuations, while another could be advantaged, suggesting a host of complex competitive pricing situations.

In sum, pricing takes on some very interesting competitive twists when costs are in one currency, and revenues in another. Theses twists become highly complex, not to say game theoretic, when firms from different home countries find themselves competing in third-country markets.

Bibliography

Cavusgil, S.T., Chan, K., and Zhang, C. (2003) Strategic orientation in export pricing: a clustering approach to create firm taxonomies. *Journal of International Marketing*, **11**, 47–72.

Clark, T.A., Kotabe, M., and Rajaratnam, D. (1999) Exchange rate pass-through and international pricing strategy: a conceptual framework and research propositions. *Journal of International Business Studies*, **30**, 249–268.

Das, N., Quelch, J., and Swartz, G. (2000) Prepare your company for global pricing. *Sloan Management Review*, **42**, 61–70.

Forman, H. and Hunt, H.J. (2005) Managing the influence of internal and external determinants on international industrial pricing strategies. *Industrial Marketing Management*, **34**, 133–146.

Myers, M.B., Cavusgil, S.T., and Diamantopoulos, A. (2002) Antecedants and actions of export pricing strategy. *European Journal of Marketing*, **36**, 159–188.

Nagle, T.T. and Holden, R.K. (1995) *The Strategy and Tactics of Pricing*, Prentice-Hall, Englewood Cliffs.

Solberg, C.A., Stottinger, B., and Yaprak, A. (2006) A Taxonomy of the pricing practices of exporting firms: evidence from Austria, Norway, and the United States. *Journal of International Marketing*, **14**, 23–48.

Stottinger, B. (2001) Strategic export pricing: a long and winding road. *Journal of International Marketing*, **9**, 40–63.

Theodosiou, M. and Katsikeas, C.S. (2001) Factors influencing the degree of international pricing strategy standardization of multinational corporations. *Journal of International Marketing*, **9**, 1–18.

countertrade

Aspy P. Palia

Many countries mandate, encourage, or permit use of countertrade – the linked trade obligations of two or more enterprises in two or more countries as stipulated in their import and export commitments. Exporters (*see* STRATEGIC EXPORT MARKETING–ACHIEVING SUCCESS IN A HARSH ENVIRONMENT) are often required to accept commodities/services in payment for the goods they sell overseas. If they accept these terms and conclude a countertrade agreement, they acquire a large inventory of countertraded commodities. Otherwise, they forego additional sales, MARKET SHARE, and entry (*see* MARKET ENTRY AND EXPANSION) into growing markets (*see* EMERGING MARKETS). Trading houses, multinational corporations (MNCs), and other countertrade specialists help dispose the surplus countertraded commodities. They specialize by services rendered, by product category, and by geographic region served (Palia, 1990).

The countertrade literature refers to a variety of transactions, which include barter, buybacks, clearing agreements, compensation, counterpurchase, evidence accounts, offsets, parallel trading, and switching (Aggarwal, 1989; Carter and Gagne, 1988; Cho, 1987; Choudhry, McGeady, and Stiff, 1989; Hennart, 1990; Huszagh and Huszagh, 1986; Khoury, 1984; Lecraw, 1989; Mirus and Yeung, 1986; Neale and Shipley, 1988; Okoroafo, 1988; Riesman, Aggarwal, and Fuh, 1989; Shipley and Neale, 1987). The US International Trade Commission recognizes five basic types of transactions: counterpurchase, compensation, offsets, barter, and switch (DeMarines, 1982).

Barter is the contractual, direct exchange of goods or services between two principals without the use of currencies. The contracting parties decide the values of the products/services to be exchanged. Barter agreements are often concluded between governments within one year. Otherwise, provisions are made to adjust the exchange ratio in response to world price fluctuations. Goods transfer is usually accomplished through a single contract. Barter transactions among governments are often accomplished through clearing agreements.

A clearing agreement between two nations specifies the goods to be exchanged, the exchange ratio, and the period during which the exchange is to be concluded. The exchange ratio is agreed to in advance and stipulated in the barter agreement. Accounts are maintained on the two-way flow of goods. At the end of the period (usually one year), any imbalance is corrected through hard-currency payments. If the clearing agreement is renewed/extended, imbalances in the flow of goods may be corrected through issuance of fresh credit.

Compensation agreements, referred to as *buyback* or *take-back*, entail the sale of plant, equipment, and/or technology. They involve the sale of a turnkey project/facility and specify that payment will be made via the plant output. Compensation transactions are generally of a higher value than other countertrade transactions. The product take-back period ranges from 5 to 20 years.

Exporters sometimes discover during the negotiations that in order for buyers in developing countries (DCs) and nonmarket economies (NMEs) to agree to the purchase contracts, the seller must agree to buy or market DC/NME products of usually less value than the original purchase contract. If the DC/NME offers products that are unrelated to the products sold by the exporter, the arrangement is referred to as *counterpurchase*. Counterpurchase transactions are relatively short term (one to five years). The counterpurchase requirement is agreed upon contractually. The seller must fulfill its contract within a specified period, failing which an agreed-upon penalty is imposed.

The offset agreement is used for defense-related contracts, commercial aircraft sales, and other priority items by the buyer government. The principals are a supplier (of defense-related equipment) in a developed country and a foreign government buyer. The supplier is required by the buyer government to assist in, or to arrange for, the marketing of products produced by firms located in the buyer country. Alternatively, the supplier must permit some portion of the exported product to be manufactured by producers in the buyer country.

Exporters that are unable to dispose the countertraded goods they receive, may decide to

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

sell them to a trading house specializing in switch trade. Switch may involve a series of complex transactions with sizable discounts before a hard-currency buyer is found. Switch-trader networks of firms and individuals offer ready markets for discounted countertraded products.

Exporters can take advantage (*see* COMPETITIVE ADVANTAGE: ITS SOURCES AND THE SEARCH FOR VALUE) of emerging opportunities in overseas markets if they are familiar with (i) the forms of countertrade practiced and (ii) the variety of trade, trade finance, consulting, international banking, marketing research, information, and other services provided by countertrade service organizations for specific products in specific markets (Palia, 1990).

Bibliography

Aggarwal, R. (1989) International business through barter and countertrade. *Long Range Planning*, **22**, 75–81.

Carter, J.R. and Gagne, J. (1988) The dos and don'ts of international countertrade. *Sloan Management Review*, **29**, 31–37.

Cho, K.R. (1987) Using countertrade as a competitive management tool. *Management International Review*, **27**, 50–57.

Choudhry, Y.A., McGeady, M., and Stiff, R. (1989) An analysis of attitudes of US firms towards countertrade. *Columbia Journal of World Business*, **24**, 31–38.

DeMarines, R.J. (1982) Analysis of Recent Trends in U.S. Countertrade, U.S. International Trade Commission, Washington, DC.

Hennart, J. (1990) Some empirical dimensions of countertrade. *Journal of International Business Studies*, **21**, 243–270.

Huszagh, S.M. and Huszagh, F.W. (1986) International barter and countertrade. *International Marketing Review*, **3**, 17–19.

Khoury, S.J. (1984) Countertrade: forms, motives, pitfalls and negotiation requisites. *Journal of Business Research*, **12**, 257–270.

Lecraw, D.J. (1989) The management of countertrade: factors influencing success. *Journal of International Business Studies*, **20**, 41–59.

Mirus, R. and Yeung, B. (1986) Economic incentives for countertrade. *Journal of International Business Studies*, **17**, 27–39.

Neale, C.W. and Shipley, D. (1988) Effects of countertrade – divergent perceptions between participants and non-participants. *Journal of Management Studies*, **25**, 57–71.

Okoroafo, S.C. (1988) Determinants of LDC mandated countertrade. *International Marketing Review*, **5**, 16–24.

Palia, A.P. (1990) Worldwide network of countertrade services. *Industrial Marketing Management*, **19**, 69–76.

Riesman, A., Aggarwal, R., and Fuh, D. (1989) Seeking out profitable countertrade opportunities. *Industrial Marketing Management*, **18**, 65–71.

Shipley, D.D. and Neale, C.W. (1987) Industrial barter and countertrade. *Industrial Marketing Management*, **16**, 1–8.

emerging markets

Chris Styles and Ranjit Voola

The term *emerging markets* was coined in 1981, and has been used to describe countries that are on a trajectory to becoming developed or advanced markets. Emerging markets are characterized by a growing middle class, the introduction of economic reforms aimed at opening up the market (including greater transparency and accountability), and a steep (above average) economic growth rate. This is in addition to a mind-set of wanting to become active participants in a global economy (*see* FORCES AFFECTING GLOBAL INTEGRATION AND GLOBAL MARKETING).

There are a number of alternative operational definitions leading to some differences in the list of countries that are classified as emerging. For example, as of January 2009, the *MSCI* (Morgan Stanley Capital Investment) *Emerging Markets Index*, based on an analysis of equity markets, consisted of the following 23 market country indices: Argentina, Brazil, Chile, China, Colombia, Czech Republic, Egypt, Hungary, India, Indonesia, Israel, Korea, Malaysia, Mexico, Morocco, Peru, Philippines, Poland, Russia, South Africa, Taiwan, Thailand, and Turkey (MSCI Barra, 2009). The importance and rapid economic growth of these markets is demonstrated by the increase in weight of the emerging markets in the MSCI All Country World Index from 1% in 1998 to 12% in 2008. Further, between December 1998 and March 2008, the value of emerging equity markets had increased by 400% when compared to developed countries' growth of 44% over the same period (Briand and Fachinotti, 2008).

While there are many definitions for emerging markets, all of these tend to have three characteristics in common: (i) rapid economic growth; (ii) proactive adoption of market-based frameworks; and (iii) being in a transition phase between developing- and developed-country status.

Markets within the emerging category have been further classified on the basis of national income and progress on market infrastructure development. For example, the *FTSE* (Financial Times and the London Stock Exchange) index divides emerging markets into advanced emerging markets (e.g., Brazil, South Korea), second emerging markets (e.g., India and Russia), and frontier markets (e.g., Bahrain and Sri Lanka). The overall methodology is based on economic size, wealth, quality of markets, and depth and breadth of the markets (FTSE, 2008).

The so-called BRIC countries (Brazil, Russia, India, and China) have been highlighted as exemplars of emerging markets. BRIC countries comprise 40% of the world population, a quarter of the earth's land coverage, and have a combined gross domestic product (GDP, purchasing power parity) in excess of 15 trillion US dollars. It is projected that by 2050 the BRIC economies have the potential to overtake the present richest countries in terms of economic wealth (Wilson and Purushothaman, 2003). A more detailed examination of India and China is instructive. India opened its economy in 1991 and its growth rate (GDP) was 9% in 2007, although this slowed to 5.3% in 2008 because of the global financial crisis. Similarly, China liberalized its economy in 1978 and had a growth rate of 13% in 2007, which again decreased to 6.8% in 2008, still considerably higher than that of developed markets (Zhixin, 2009). In addition, both countries have developed internal economic conditions and regulations that encourage local multinational companies (such as Tata in India and Lenovo in China) to actively compete with multinationals in advanced markets (*see* GLOBAL MARKETING STRATEGY; GLOBAL PRODUCT R&D; GLOBAL SOURCING STRATEGY: AN EVOLUTION). Both India and China also have a growing middle class – the World Bank estimates that by 2030 developing countries will account for 93% of the middle class globally, with China and India representing two-thirds of this growth (China 52%; India 12% increase) (Knowledge@Wharton, 2008).

The key challenges facing emerging markets include governance, political risk, rapid urbanization, a shortage of highly skilled labor, poor infrastructure, and more recently, the global economic slowdown that began in 2008. From a strategic point of view, the challenge for marketers is to identify viable segments with A FRAMEWORK FOR CREATING VALUE PROPOSITIONS relevant to local customers

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

(*see* MARKETING STRATEGY IMPLEMENTATION). Finding the correct price point (*see* INTERNATIONAL PRICING OBJECTIVES AND STRATEGIES) is also important to ensure affordability. Infrastructure, regulations, and local networks that can be difficult to break into are key strategic implementation issues (*see* MARKETING STRATEGY IMPLEMENTATION). The key is often to find a strong, well-connected local partner (*see* INTERNATIONAL MARKETING CHANNELS).

Bibliography

Briand, R. and Fachinotti, G. (2008) Emerging Markets: A 20 Year Perspective, MSCI Barra available at: http://www.mscibarra.com/products/indices/em_20/EM_20_Anniversary.pdf (accessed 2009).

FTSE (2008) FTSE Global Equity Index Series Country Classification, available at: http://www.ftse.com/Indices/Country_Classification/Downloads/FTSE_Country_Classification_Sept_08_update.pdf (accessed September 2008, update).

Knowledge@Wharton (2008) The New Global Middle Class: Potentially Profitable – but also Unpredictable, available at: http://knowledge.wharton.upenn.edu/article.cfm?articleid=2011# (accessed 2009).

MSCI Barra (2009) MSCI Emerging Markets, available at: http://www.mscibarra.com/products/indices/licd/em.html (accessed 2009).

Wilson, D. and Purushothaman, R. (2003) Dreaming with BRICs: The Path to 2050, Global economics paper No: 99, Goldman Sachs.

Zhixin, D. (2009) GDP Growth Slows to 6.8% in Q4 of 2008, China Daily, Available at: http://www.chinadaily.cn/china/2009-01/22/content_7420728.htm (accessed 22 January 2009).

global sales management

Earl D. Honeycutt, Jr.

The sales-related activities that firms and organizations engage in when operating across borders in more than one country (*see* GLOBAL MARKETING STRATEGY). Sales executives and managers must master the challenges of working across national and cultural borders that exist in a global economy (*see* GLOBAL ACCOUNT MANAGEMENT: THE RATIONALE AND MOTIVATION). Global sales management can be divided into sales strategies (*see* SALES FORCE STRATEGY) and human resource practices.

Firms that sell their products overseas do so on a continuum from indirect partnerships to direct participation. On the basis of individual market situation, managers must decide upon the numbers, characteristics, and assignments of sales personnel (*see* INTERNATIONAL MARKETING CHANNELS). For example, companies entering the global market can find partners who provide existing salespersons or sales agents that understand the local marketplace, the local culture, and language (*see* SOCIETY, CULTURE, AND GLOBAL CONSUMER CULTURE) or salespersons from the home country can be assigned abroad. Firms are moving from multinational to transnational status. Transnational firms see a borderless global market with myriad locations to acquire natural resources, site production facilities, hire operations and marketing talent, offer high-quality products and services, and develop a global brand image. Multinational and transnational firms use three types of salespersons and sales managers: expatriates, local nationals, and third-country nationals. Each category of salesperson has advantages and disadvantages. Expatriates are sent from the country of the firm's headquarters and possess high company and product knowledge. However, they may be weak in the local culture and language. Home-country sales personnel understand local markets and cultures, but may be less dedicated to the firm and have less influence with headquarters. Third-country nationals originate from countries outside the work nation or firm country, speak several languages, and understand local markets and culture. Recently, as a result of cost and expertise, companies are relying more on local managers to represent the global firm.

In the area of human resources, global sales managers must understand the lens through which they view the behavior of overseas sales representatives and sales managers. Managers who view sales situations and sales force behavior based upon how things are done in the home country, and do not understand local cultural practices, are likely to make serious "*faux pas*" or social blunders due to ethnocentrism (*see* CONSUMER ACCULTURATION). Poor decisions include selecting the wrong salesperson due to applicants not responding to questions/situations the way the sales manager expects; providing inappropriate training – it is easier to standardize technical knowledge and company information, but important to localize the selling process and competitive situation. Global sales personnel are also less likely to be motivated by money alone. For example, in Japan, the focus is on corporate culture and group interaction to motivate sales personnel. Lastly, evaluation of global sales personnel is influenced by the salesperson's level of individualism and perceptions of environmental control. To improve success, transnational firms select their global sales managers early and provide them with expertise and cultural intelligence training so that they can successfully manage workers from diverse backgrounds and cultures.

Bibliography

Bartlett, C.A. and Ghoshal, S. (2003) What is a global manager? *Harvard Business Review*, August, **81** (8), 101–108.

Early, C. and Mosakowski, E. (2004) Cultural intelligence. *Harvard Business Review*, October, **82** (10), 139–146.

Honeycutt, E.D., Ford, J.B., and Simintiras, A. (2003) *Sales Management: A Global Perspective*, Routledge, London.

Macquin, A., Rouzies, D., and Prime, N. (2000) The influence of culture on personal selling interactions. *Journal of Euromarketing*, **9** (2), 71–88.

Shay, J. and Baack, S.A. (2004) Expatriate assignment, adjustment and effectiveness: an empirical examination of the big picture. *Journal of International Business Studies*, **35** (3), 216–232.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

international trade intermediaries

Paul D. Ellis

International trade intermediaries (ITIs) are specialist intermediaries that are involved in the distribution of goods across national borders. ITIs are sometimes known as *trading companies*, *export management companies*, *sogo shosha*, or simply, *importer–exporters*. ITIs may be involved in many activities and perform a variety of services, but their core activity is the management of international trade. This service either involves an agency-brokerage role (selling on behalf of another) or a merchant-reselling (taking title to the goods traded) role (Casson, 1998). ITIs are ubiquitous in global markets and are one of the most enduring organizational entities in the history of international business. Historically, ITIs have played a prominent role in the development of international trade and marketing in many societies (Ellis, 2003a; Carlos and Nicholas, 1988).

Acting as agents, ITIs exist because manufacturers sometimes lack the ability or the incentive to engage in exporting directly (*see* STRATEGIC EXPORT MARKETING–ACHIEVING SUCCESS IN A HARSH ENVIRONMENT). By consolidating orders from multiple suppliers, building trade-supporting infrastructure, and developing logistical expertise, ITIs can improve the efficiency of distribution channels lowering the costs of exchange for others. In return, the intermediary typically charges a commission based on the cost of goods sold.

Manufacturers may hesitate to contract with independent intermediaries because of the threat of opportunism. This threat is central to the agency dilemma and arises because the manufacturer and intermediary possess different information about the exporting task. This imbalance or asymmetry may be exploited by the ITI. For example, to win a contract, the ITI may be tempted to conceal or misrepresent information regarding their abilities and resources. If the manufacturer as the principal cannot fully verify these claims in advance, they may suffer the consequences of adverse selection. Another potential problem, moral hazard, may arise after the contract has been signed when the principal is unable to ensure that the ITI as the agent delivers the agreed-upon level of effort.

The hazards of relying on opportunistic intermediaries may render the agency relationship unattractive to manufacturers compelling them to pursue other modes of exchange, such as direct exporting. The circumstances under which manufacturers would prefer ITIs to in-house exporting can be predicted using transaction cost analysis, which describes the matching of exchange characteristics with particular exchange modes or governance structures (Anderson and Gatignon, 1986; Klein, Frazier, and Roth, 1990). For example, transaction cost analysis would predict a preference for direct or integrated channels (*see* INTERNATIONAL MARKETING CHANNELS) when traded products are complex and require a high level of after-sales service. In this type of exchange, the costs of performing intermediary functions in house are likely to be less than the combined ex ante costs of negotiating stringent contracts outlining the expected level of service and the ex post costs associated with monitoring and enforcing contractual obligations.

Acting as merchant resellers, ITIs create new markets for goods and services by developing exchanges between buyers and sellers located in different nations. In developing these links, the ITI may also transfer marketing skills to others in the channel. This may happen intentionally, as when an intermediary provides designs or materials for a new product, or unintentionally in the routine execution of the marketing function.

Although ITIs improve the efficiency of exchange and entrepreneurially create markets where none existed previously, their survival within cross-border channels of distribution is inherently impermanent. The information asymmetries that give ITIs their "edge" in a market tend to diminish over time undermining their position in the channel. As manufacturers gain experience dealing with foreign buyers, they may consider integrating intermediary functions and exporting directly to foreign markets. Indeed, by exceeding their client's performance expectations, the intermediary may even hasten their own removal by signaling the benefits of going direct. This gives rise to the traders' dilemma, which describes the increased risk of replacement attached to extreme levels of

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

performance: Poor-performing intermediaries risk being replaced by other intermediaries, while high-performing intermediaries may be eliminated by upstream suppliers or downstream buyers keen to reap the benefits of more direct exchange modes (Ellis, 2005). In view of these threats to their survival, intermediaries continually seek out new opportunities to broker relationships. Consequently, the survival of the ITI is a function of (i) prolonging the inevitable decline of existing exchange relationships and (ii) replacing lost business with new relationships. When the addition of income from new business no longer matches the loss of income from exiting clients, the intermediary will have started down the path to organizational decline.

Bibliography

Anderson, E. and Gatignon, H. (1986) Modes of foreign entry: a transaction cost analysis and propositions. *Journal of International Business Studies*, **17** (3), 1–26.

Bello, D.C. and Williamson, N.C. (1985) Contractual arrangement and marketing practices in the indirect export channel. *Journal of International Business Studies*, **16**, 65–82.

Carlos, A.M. and Nicholas, S. (1988) Giants of an earlier capitalism: the chartered trading companies as modern multinationals. *Business History Review*, **62** (Autumn), 398–419.

Casson M. (1998) The economic analysis of multinational trading companies, in *The Multinational Traders* (ed. G. Jones), Routledge, London, pp. 22–47.

Ellis, P.D. (2003a) Are international trade intermediaries catalysts in economic development? A new research agenda. *Journal of International Marketing*, **11** (1), 73–93.

Ellis, P.D. (2003b) Social structure and intermediation: market-making strategies in international exchange. *Journal of Management Studies*, **40** (7), 1677–1702.

Ellis, P.D. (2005) The traders' dilemma: the adverse consequences of superior performance in mediated exchanges. *International Business Review*, **14** (4), 375–396.

Klein, S., Frazier, G., and Roth, V. (1990) A transaction cost analysis model of channel integration in international markets. *Journal of Marketing Research*, **26**, 196–208.

Peng, M.W. and Ilinitch, A.Y. (1998) Export intermediary firms: a note on export development research. *Journal of International Business Studies*, **29** (3), 609–620.

international negotiations

Bruce Money

Negotiating deal points is fundamental to any business transaction. Basic negotiation issues have come to be widely studied, such as whether one should anchor the transaction with a first offer, how negotiators respond to the pressures of a deadline, and decisions about how to make sensible concessions (*see* CONSUMER DECISION MAKING). Classic models of negotiation include the problem-solving approach (Graham, 1987), the relationships, parties' behaviors, and influencing conditions (RBC) model (Weiss, 1993), and the branching-chain model of multilateral negotiations (Pruitt, 1994).

It follows that conducting effective cross-national and/or cross-cultural negotiations is becoming increasingly important in today's global marketplace (*see* GLOBAL MARKETING STRATEGY). International business negotiations have their own uniqueness. Overcoming the obvious challenges of language differences and distant locations are just the beginning of effective negotiations with overseas clients. More subtle issues are numerous, and invoke questions such as the following

- What do people in other countries consider good negotiation?
- How can a negotiator be effective in a culture that is different from his or her or own? (*see* CROSS-CULTURAL PSYCHOLOGY OF CONSUMER BEHAVIOR)
- How do references to time, power, and other tactics differ in international contexts?
- What other global factors affect negotiation?

International negotiation presupposes an overlay of national culture and other cross-national factors on any of these models. According to Reynolds, Simintiras and Vlachou (2003), noncultural factors may include environmental and organizational factors (e.g., political or legal issues), individual negotiator characteristics (such as communication effectiveness or negotiation experience), factors related to the negotiation itself (such as norms and objectives), and negotiation outcomes (objective and subjective measures of success). This framework loosely aligns that of Graham's (1987) model of personal *bargainer characteristics* (such as listening ability; combined with *situational factors* (e.g., physical setting,) to influence the *process* of negotiation (integrative vs. distributive; PERSUASION, to affect the *outcome* (satisfaction).

National culture probably looms the largest of the factors affecting international negotiation, the subject of dozens of studies (*see* BASE OF THE PYRAMID MARKETS: CULTURE INSIGHTS AND MARKETING IMPLICATIONS). For example, in Graham's model, culture would be one of the personal characteristics, and culture differences a situational factor. Two of the most prominent frameworks of national culture incorporated by most cross-national negotiation researchers are Hofstede's (2001) five dimensions of culture and the construct of context by Hall and Hall (1990). Briefly, Hofstede's dimensions of individualism/collectivism, power distance, uncertainty avoidance, masculinity, and Confucian dynamism, have been found to affect the negotiation process, depending on the dimension and negotiation phenomenon, for example, multilateral negotiations (Money, 1998).

Here mention should be made of the well-placed debate regarding the applicability of Hofstede measures at the individual behavior (Spector, Cooper, and Sparks, 2001), which has come to be known as the *ecological fallacy*. Values are more fine grained than index scores and need to be measured at the individual level, while recognizing the validity limitations of the Hofstede measures (Bearden, Money, and Nevins, 2006) (*see* VALIDITY IN EXPERIMENTATION).

The other cultural construct important to negotiation behavior is that of context, a measure of the implicit "code" in communication. Cultures are characterized as either high or low context (Hall and Hall, 1990) in their communication. In high-context cultures, the social and temporal setting of communications is key to understanding the intended meaning of messages, that is, the literal content of a message may be less important than the situation by which the message is conveyed. Low-context cultures, in contrast, do not depend as much on contextual factors in communication; rather the

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

words of a message are interpreted more literally. In other words, context is the amount of implicit communication "code" in a culture. Individuals from high-context cultures, such as those in Asia or Latin America, value face-saving, innuendo, and "between the lines" communication. More importantly, they are often misunderstood in negotiations by counterparts from low-context cultures such as the United States or Germany, where more explicit communications place emphasis on the message content (Weiss, 1993).

Finally, a recent review of international negotiation phenomena (Brett, 2007) suggests that other international factors may affect negotiation processes and outcomes (*see* GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES), for example, composition of multicultural teams in organizations internally, and externally, societal dilemmas, and government policies (*see* PRODUCT ETHNICITY) that may face and involve corporations and other organizations.

Bibliography

Bearden, W.O., Money, R.B. and Nevins, J.L. (2006) Assessing individual-level differences in national vulture values: evaluation of the Hofstede measures. *Journal of Business Research*, **59**, 195–203.

Brett, J.M. (2007) *Negotiating Globally*, Jossey-Bass, San Francisco, CA.

Graham, J.L. (1987) A theory of interorganizational negotiations. *Research in Marketing*, **9**, 163–183.

Hall, E.T. and Hall, M.R. (1990) *Understanding Cultural Differences*, Intercultural Press, Yarmouth, ME.

Hofstede, G. (2001) *Culture's Consequences*, Sage, Beverly Hills, CA.

Money, R.B. (1998) International multilateral negotiations and social networks. *Journal of International Business Studies*, **29**, 695–710.

Pruitt, D.G. (1994) Negotiation between organizations: a branching chain model. *Negotiation Journal*, **10**, 217–229.

Reynolds, N., Simintiras, A. and Vlachou, E. (2003) International business negotiations, present knowledge and direction for future research. *International Marketing Review*, **20**, 236–261.

Spector, P.E., Cooper, C.L. and Sparks, K. (2001) An international study of the psychometric properties of the Hofstede values survey module of 1994: a comparison of individual and country/province level results. *Applied Psychology: An International Review*, **50**, 269–281.

Weiss, S.E. (1993) Analysis of complex negotiations in international business: The RBC perspective. *Organization Science*, **4**, 269–300.

international product diffusion

Eden Yin

International product diffusion is becoming an increasingly important topic for both academics and industry practitioners as rapid globalization expedites market integration, which enables more goods and services to be diffused internationally. There is a growing literature on this topic. Most of these studies examine the differences in the diffusion process across nations and factors explaining the differences.

The general finding from these studies is that a nation's wealth, culture, social factors, and the time of its introduction are the critical determinants shaping a new product's diffusion process across countries. As for the wealth factor, even though an early study does not find it important in driving the cross-country takeoff (Tellis, Stremersch, and Yin, 2003), other studies prove otherwise (Chandrasekaran and Tellis, 2008). Moreover, a country's wealth is believed to affect the growth rate and duration of a new product across nations (Stremersch and Tellis, 2004). At the same time, income distribution and the level of international trade is proven to influence the penetration level, adoption rate, and speed of a product's diffusion (Talukdar, Sudhir, and Ainslie 2002). Such a conclusion also applies to the international diffusion of technology products subject to network externalities (Dekimpe, Parker, and Sarvary, 2000).

Research also finds culture, supposedly the most prominent factor in shaping international product diffusion, plays either a significant or partially significant role in affecting this process. Apparently, high context culture leads to faster adoption (Takada and Jain, 1991). At the same time, culture also shapes both growth rate and duration of a new product across countries even though to a lesser extent than economic wealth (Stremersch and Tellis, 2004). Moreover, cultural traits such as higher need for achievement, higher level of collectivism, industriousness, and lower uncertainty avoidance are associated with a quicker takeoff (Chandrasekaran and Tellis, 2008), or explain only intercountry differences in time-to-takeoff partially (Tellis, Stremersch, and Yin, 2003).

As for social factors, a number of studies indicate that the percentage of women in labor force (Talukdar, Sudhir, and Ainslie, 2002, Gatignon, Eliashberg, and Robertson, 1989), mobility (Gatignon, Eliashberg, and Robertson, 1989), cosmopolitanism or level of urbanization (Gatignon, Eliashberg, and Robertson, 1989, Talukdar, Sudhir, and Ainslie, 2002), heterogeneity of the country's social system (Dekimpe, Parker, and Sarvary, 2000), type of communication (Takada and Jain, 1991, Putsis *et al.*, 1997), and so on significantly affect the process of international product diffusion. In particular, Takada and Jain (1991) conclude that the rate of adoption in countries characterized by homophilous communication is faster than that in countries characterized by heterophilous communication, while Putsis *et al.* (1997) suggest that focusing on the most gregarious countries will maximize adoption in subsequent countries.

Studies also identify other sets of factors that influence the diffusion process of products in international markets. For instance, the time of introduction and therefore the introductory lag affect the penetration level, adoption rate, and speed of product diffusion (Talukdar, Sudhir, and Ainslie, 2002). The later a product is introduced in a market, the faster will be the rate of adoption (Takada and Jain, 1991). With regard to the product takeoff, besides culture and wealth, product class, product vintage, and prior takeoff are also important drivers for it to occur (Chandrasekaran and Tellis, 2008). In the context of technology products, installed base and international experience with this technology also affect the diffusion process (Dekimpe, Parker, and Sarvary, 2000). The direct managerial implication from these studies is whether a firm should adopt a sprinkler, that is, simultaneous introduction, or waterfall strategies, that is, sequential introduction, when they launch their new product globally. Some researchers support the former (Ohmae, 1985), while others endorse the latter (Putsis *et al.*, 1997, Tellis, Stremersch, and Yin, 2003; Stremersch and Tellis, 2004). In fact, the tradeoff between a waterfall and sprinkler strategy reduces to a tradeoff between sales maximization, for example, sprinkler, and risk minimization, for example, waterfall.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

See also *consumer innovativeness; innovation diffusion; global marketing strategy*

Bibliography

Chandrasekaran, D. and Tellis, G.J. (2008) Global takeoff of new products: culture, wealth, or vanishing differences? *Marketing Science*, **27** (5), 844–860.

Dekimpe, M.G., Parker, P.M. and Sarvary, M. (2000) Global diffusion of technological innovations: a coupled-hazard approach. *Journal of Marketing Research*, **XXXVII**, 47–59.

Gatignon, H., Eliashberg, H. and Robertson, T.S. (1989) Modeling multinational diffusion patterns: an efficient methodology. *Marketing Science*, **8** (3), 231–247.

Ohmae, K. (1985) *Triad Power*, The Free Press, New York.

Putsis, W.P., Balasubramanian, S., Kaplan, E.H. and Sen, S.K. (1997) Mixing behavior in cross-country diffusion, *Marketing Science*, **16** (4), 354–369.

Stremersch, S. and Tellis, G.J. (2004) Understanding and managing international growth of new products. *International Journal of Research in Marketing*, **21** (4), 421–438.

Takada, H. and Jain, D. (1991) Cross-national analysis of diffusion of consumer durable goods in pacific rim countries2. *Journal of Marketing*, **55**, 48–54.

Talukdar, D., Sudhir, K. and Ainslie, A. (2002) Investigating new product diffusion across products and countries. *Marketing Science*, **21** (1), 97–114.

Tellis, G.J., Stremersch, S. and Yin, E. (2003) The international takeoff of new products: the role of economics, culture, and country innovativeness. *Marketing Science*, **22** (2), 188–208.

marketing aspects of psychic distance

Jody Evans

The term *psychic distance* was first introduced by Beckerman (1956) in the concluding paragraph of his study on the distribution of international trade. It is the work done at the Uppsala University during the 1970s, however, that is credited with developing and popularizing the construct. Johanson and Wiedersheim-Paul (1975, p. 308) defined *psychic distance* as those "factors preventing or disturbing flows of information between firm and market." O'Grady and Lane (1996) redefined *psychic distance* as "... a firm's degree of uncertainty about a foreign market resulting from cultural differences and other business difficulties that present barriers to learning about the market and operating there" (O'Grady and Lane, 1996, p. 330). More recently, Evans and Mavondo (2002, p. 517) focused on the aspect of perception and defined *psychic distance* as "the distance between the home market and a foreign market, resulting from the perception of both cultural and business differences."

While the definition has evolved over time, there is a general consensus that psychic distance relates to differences between the home and foreign market that impede learning and foster uncertainty. There is also significant overlap in the lists of constituent elements. These include differences in culture, language, religion, education, legislation, politics, economic conditions, market structure, and business practices. The primary source of current debate relates to the manner in which psychic distance is operationalized. There are four debatable issues: (i) perceived versus objective measures; (ii) national versus individual measures; (iii) cultural distance (*see* MARKETING ASPECTS OF CULTURAL DISTANCE) versus psychic distance, and (iv) the relative importance of contributing factors (Dow and Karunaratna, 2006).

In terms of outcomes, psychic distance has been linked to foreign market selection, entry mode (*see* MARKET ENTRY AND EXPANSION), the standardization/adaptation of marketing strategies (*see* STANDARDIZATION/ADAPTATION OF INTERNATIONAL MARKETING STRATEGY), and firm performance (*see* EXPORT PERFORMANCE). Psychic distance has traditionally been associated with the internationalization process. In particular, the underlying premise of the Uppsala model is that the internationalization of a firm is an incremental process where firms initially enter markets that are psychically similar and successively expand into more distant foreign markets. This positive association between psychic distance and the sequence of market entry has received limited support. There are suggestions that in the dynamic global marketplace, other factors may play a more important role in determining market selection or may confound the nature of the psychic distance–internationalization process relationship (Ellis, 2008).

Psychic distance plays a critical role in both entry-mode (*see* MARKET ENTRY AND EXPANSION) and marketing-strategy decisions (*see* STANDARDIZATION/ADAPTATION OF INTERNATIONAL MARKETING STRATEGY). The greater the psychic distance between the home market and a foreign market, the more likely a firm is to adopt an entry mode that requires less financial commitment and lower levels of sole control. This can be attributed to the fact that firms will perceive greater risk in markets that are psychically distant and, consequently, will seek to minimize their financial exposure. In terms of marketing strategy, the feasibility of a predominantly standardized strategy is driven by market similarity. Correspondingly, psychic distance is associated with adaptation of marketing strategy where firms may choose (discretionary) to adapt their strategy or are forced (obligatory) to respond to market differences.

A paradoxical relationship is said to exist in regard to psychic distance and firm performance (*see* EXPORT PERFORMANCE). O'Grady and Lane (1996) suggest that an assumption of similarity prevents executives from noticing subtle, but important, differences in psychically close markets, which is strongly associated with poor organizational performance. Evans and Mavondo (2002) contend that when firms enter psychically distant markets the increased sense of risk and uncertainty will elicit a strong desire to learn more about the market. This knowledge

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

will enhance the firm's strategic decisions and, consequently, its performance.

Bibliography

Beckerman, W. (1956) Distance and the pattern of Intra-European trade. *Review of Economics and Statistics*, **28**, 31–40.

Dow, D. and Karunaratna, A. (2006) Developing a multidimensional instrument to measure psychic distance stimuli. *Journal of International Business Studies*, **37**, 578–602.

Ellis, P.D. (2008) Does psychic distance moderate the market size–entry sequence relationship? *Journal of International Business Studies*, **39**, 351–369.

Evans, J. and Mavondo, F. (2002) Psychic distance and organizational performance: an empirical examination of international retailing operations. *Journal of International Business Studies*, **33**, 515–532.

Johanson, J. and Wiedersheim-Paul, F. (1975) The internationalization of the firm – four Swedish cases. *Journal of Management Studies*, **12**, 305–322.

O'Grady, S. and Lane, H. (1996) The psychic distance paradox. *Journal of International Business Studies*, **27**, 309–333.

marketing aspects of cultural distance

Carlos M. P. Sousa

There is a general consensus in the literature that when firms decide to enter foreign markets, they must adjust to a different cultural environment and be prepared for challenges, such as differences in language, lifestyles, cultural standards, consumer preferences, and purchasing power, among others (*see* GLOBAL MARKETING STRATEGY; CROSS-CULTURAL PSYCHOLOGY OF CONSUMER BEHAVIOR; STANDARDIZATION/ADAPTATION OF INTERNATIONAL MARKETING STRATEGY; BASE OF THE PYRAMID MARKETS: CULTURE INSIGHTS AND MARKETING IMPLICATIONS). To assess these differences between countries, a new body of literature has emerged around the concept of cultural distance, which has been used to measure the scope and scale of these differences. The assumption is that cultural differences between the home and the foreign markets create a distance, which, in turn, influences the activity of the firm in the international arena. As a result, few concepts in the international-marketing literature have gained broader attention than "cultural distance." This concept has been identified as a key variable in explaining the behavior of firms in a vast array of areas including entry-mode choice, foreign-market selection, level of control (*see* MARKET ENTRY AND EXPANSION), international-marketing strategies (*see* GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES), and performance (*see* EXPORT PERFORMANCE). Mixed empirical results, however, have been found in the literature regarding the importance of this construct. Whereas some studies have found cultural distance to significantly influence the activity of the firm in the international arena (Barkema and Vermeulen, 1997), other studies have found no significant or clear relationship (Mitra and Golder, 2002).

Various reasons have been proposed to explain these inconclusive results. On the one hand, some scholars have argued that the preferences and tastes of consumers in different countries are converging to a global norm (Levitt, 1983), and hence the effect of cultural distance is likely to dilute progressively. The notion that national cultures are converging appears to be plausible considering the emergence of the Internet, greater ease of information flow, and more frequent/easier international travel. However, this argument is disputed by Barkema and Vermeulen (1997), who use data spanning almost three decades (1966–1994) in their study, and find that the effect of cultural distance does not decrease over time, that is, that cultural values and distances remain stable. This may be due to the fact that such changes concern convergences in superficial appearances of culture (i.e., symbols, heroes, and rituals) and that they do not necessarily signal a convergence in the values embedded in national cultures (Hofstede, 2006).

The inconsistencies could also be attributed to the difficulty that exists in the conceptualization and operationalization of the cultural-distance construct. For instance, as Sousa and Bradley (2006, 2008) note, many authors have used cultural distance and *psychic distance* (*see* MARKETING ASPECTS OF PSYCHIC DISTANCE) interchangeably with no clear distinction between them, a contamination that has been disputed by these authors and which may undermine the validity of the research (*see* RESEARCH RELIABILITY AND VALIDITY). *Cultural distance* has been defined as the degree to which cultural values in one country are different from those in another country (Sousa and Bradley, 2006). Accordingly, it should be used to assess differences in national culture. Several frameworks have been advanced and used in the measurement of the extent to which different cultures are similar or different, such as that of Kluckhohn and Strodtbeck (1961) and Ronen and Shenkar (1985). However, the most widely used framework is that of Hofstede's (1980) cultural dimensions. To arrive at a measure of cultural distance, Kogut and Singh (1988) were the first to combine Hofstede's dimensions into one aggregate measure of cultural distance among countries, and many studies have subsequently used this formula or an adapted version, as a measure of cultural distance. However, while a tangible and convenient tool, the Kogut and Singh index has been liable to the same criticism as that leveled against Hofstede's value

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

dimensions, for example, nonexhaustiveness, outdated data, assumptions about the linearity, additivity, and normal distributions of scores (Shenkar, 2001). As a result, there has been an increasing call in the literature to develop new measures (*see* VALIDITY AND RELIABILITY) to operationalize cultural distance as it may be unrealistic to expect that a single measure can fully discern the underlying differences across countries (Tihanyi, Griffith, and Russell, 2005). Some recent studies may offer some interesting alternatives to measure cultural distance (Schwartz, 1994) and should, therefore, be explored in future studies.

Bibliography

Barkema, H.G. and Vermeulen, F. (1997) What differences in the cultural backgrounds of partners are detrimental for international joint ventures. *Journal of International Business Studies*, **28** (4), 845–864.

Hofstede, G. (1980) *Culture's Consequences: International Differences in Work-related Values*, Sage Publications, Beverly Hills.

Hofstede, G. (2006) What did GLOBE really measure? Researchers' minds versus respondents' minds. *Journal of International Business Studies*, **37** (6), 882–896.

Kluckhohn, F.R. and Strodtbeck, F.L. (1961) *Variations in Value Orientations*, Row, Peterson, Evanston.

Kogut, B. and Singh, H. (1988) The effect of national culture on the choice of entry mode. *Journal of International Business Studies*, **19** (3), 411–432.

Levitt, T. (1983) The globalization of markets. *Harvard Business Review*, **61** (3), 92–102.

Mitra, D. and Golder, P.N. (2002) Whose culture matters? Near-market knowledge and its impact on foreign market entry timing. *Journal of Marketing Research*, **39** (3), 350–365.

Ronen, S. and Shenkar, O. (1985) Clustering countries on attitudinal dimensions: a review and synthesis. *Academy of Management Review*, **10** (3), 435–454.

Schwartz S.H. (1994) in *Individualism and Collectivism: Theory, Method and Applications* (eds U. Kim H.C. Triandis C. Kagitçibasi *et al.*), Sage Publications, Thousand Oaks, pp. 85–119.

Shenkar, O. (2001) Cultural distance revisited: towards a more rigorous conceptualization and measurement of cultural differences. *Journal of International Business Studies*, **32** (3), 519–535.

Sousa, C.M.P. and Bradley, F. (2006) Cultural distance and psychic distance: two peas in a pod? *Journal of International Marketing*, **14** (1), 49–70.

Sousa, C.M.P. and Bradley, F. (2008) Cultural distance and psychic distance: refinements in conceptualisation and measurement. *Journal of Marketing Management*, **24** (5/6), 467–488.

Tihanyi, L., Griffith, D.A., and Russell, C.J. (2005) The effect of cultural distance on entry mode choice, international diversification, and MNE performance: a meta-analysis. *Journal of International Business Studies*, **36** (3), 270–283.

consumer animosity

Jill Gabrielle Klein

Our world is fraught with international conflict that leads people to feel angry at other countries. This anger is sometimes manifested in the decisions consumers make about what to buy. Consumer decisions, in other words, can depend on more than just the material value of a product. They can depend upon how people feel about the country where the product was made.

The animosity model of foreign product purchase demonstrates that animosity, defined as the remnants of antipathy toward another country due to past military, political, or economic conflicts, affects the buying of foreign goods (Figure 1). Numerous studies have been conducted on consumer animosity since the introduction of the model to the literature in 1998. The context of the first study was Nanjing, China, a city that was the site of the horrific slaughter of approximately 300,000 civilians by the Japanese in December 1937 and January 1938. Survey research demonstrated that Nanjing consumers who were still angry with Japan because of the massacre were less willing to buy Japanese goods, and actually owned fewer Japanese electronic and durable products (Klein, Ettenson, and Morris, 1998). Interestingly, angry consumers did not denigrate the *quality* of Japanese products. They acknowledged the good workmanship, reliability, and technological advancement of Japanese products, but they still refused to buy them.

This fundamental premise of the animosity model – that effects of animosity on buying are direct and independent of product quality judgments – diverges not only from traditional country-of-origin research (*see* "COUNTRY OF ORIGIN" AS BRAND ELEMENT) but also from most behavioral frameworks in marketing which assume a primary relationship between consumers' product judgments and their purchase behavior (e.g., Green and Srinivasan, 1990). In the case of consumer animosity, anger can lead consumers to eschew a country's goods in spite of positive product perceptions.

Subsequent demonstrations of the effects of consumer animosity have shown wide generality of the model across the globe (see Riefler and Diamantopoulos, 2007). For example, Dutch consumers who are still angry at Germany because of World War II, are less likely to purchase German products (Nijssen and Douglas 2004). Japanese consumers who remain angry with the United States for dropping the atomic bomb on Hiroshima and Nagasaki still shun American products (Klein and DeBroux, 2001). Animosity can even affect how consumers feel about buying products from within their own country. Southerners in the United States whose anger still lingers from the Civil War prefer to purchase products produced in the south rather than the north (Shimp, Dunn, and Klein, 2004).

The animosity model also includes the construct of consumer ethnocentrism which is the tendency to believe that it is inappropriate or even immoral to purchase foreign products because doing so is harmful to the domestic economy (Shimp and Sharma, 1987). Previous studies have found that those high on the CETSCALE, which measures consumer ethnocentrism, are less willing to purchase foreign products, and view foreign products as lower in quality than domestic products (Netemeyer, Durvasula, and Lichtenstein, 1991).

Thus, while consumer ethnocentrism is related to *both* product judgments and purchase intentions, animosity affects consumers' purchase decisions *independently* of product judgments. A further distinction between the constructs is that animosity is comprised of consumer feelings toward a specific country, whereas consumer ethnocentrism concerns attitudes toward buying goods from *all* foreign countries. While some consumers may find it perfectly acceptable to buy foreign products from a variety of countries, they may refuse to buy a product from a specific nation toward which they feel enmity.

International conflict has many dire consequences, as we see on our television screens on a daily basis. From a consumer behavior standpoint, these tensions have the additional effect that consumers may avoid the products of the offending country. And just as the conflicts of decades ago still have an impact on current consumer choices, the conflicts of today may will

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

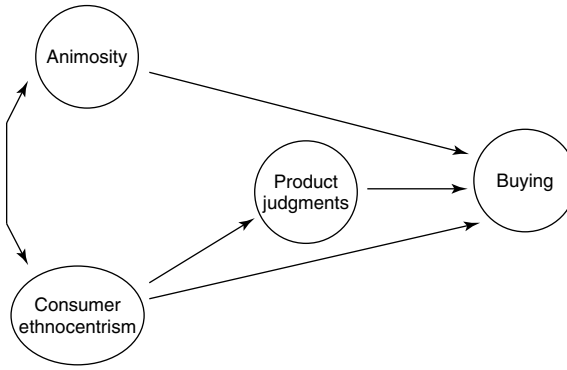


Figure 1 The animosity model of foreign product purchase.

cast their shadow on consumption decisions for decades to come.

Bibliography

Green, P.E. and Srinivasan, V. (1990) Conjoint analysis in marketing: new developments with implications for research and practice. *Journal of Marketing*, **54**, 3–19.

Klein, J.G. and DeBroux, P. (2001) Subjective norms and the animosity model: social pressure within a collectivist culture, in *European Advances in Consumer Research*, Vol. **5** (eds A. Groeppel-Klein and F.-R. Esch). Association for Consumer Research, Provo, pp. 120–123.

Klein, J.G., Ettenson, R.E., and Morris, M.D. (1998) The animosity model of foreign product purchase: an empirical test in the people's Republic of China. *Journal of Marketing*, **62**, 89–100.

Netemeyer, R., Durvasula, S., and Lichtenstein, D. (1991) A cross-national assessment of the reliability and validity of the CETSCALE. *Journal of Marketing Research*, **28**, 320–327.

Nijssen, E.J. and Douglas, S.P. (2004) Examining the animosity model in a country with a high level of foreign trade. *International Journal of Research in Marketing*, **21** (1), 23–38.

Riefler, P. and Diamantopoulos, A. (2007) Consumer animosity: a literature review and a reconsideration of its measurement. *International Marketing Review*, **24** (1), 87–119.

Shimp, T., Dunn, T., and Klein, J.G. (2004) Remnants of the U.S. Civil War and modern consumer behavior. *Psychology and Marketing*, **21** (2), 75–91.

Shimp, T. and Sharma, S. (1987) Consumer ethnocentrism: construction and validation of the CETSCALE. *Journal of Marketing Research*, **24**, 280–289.

consumer affinity construct

Adamantios Diamantopoulos and Eva M. Oberecker

Sentiments toward foreign countries as captured by constructs such as CONSUMER ANIMOSITY, consumer ethnocentrisim, consumer cosmopolitanism[1] or xenophilia,[2] have potential power in explaining consumption behavior toward foreign products (*see* PRODUCT ETHNICITY) and are thus of importance to international marketers. The consumer affinity construct specifically captures *favorable* sentiments toward a *specific* foreign country, which "affect[s] behavioral consequences (i.e., conations), such as intentions to consume products, brands, and services from the affinity country" (Oberecker, Riefler, and Diamantopoulos, 2008, p. 25).

The conceptual core of the affinity construct is *positive affect*[3] and drawing from social identity theory (Tajfel, 1982), affinity builds on the distinction between in- and out-groups and assumes that an out-group can potentially be included into one's in-group(s). In a country context, this means that due to affinity for a particular foreign country, the latter foreign country may supplement the in-group (i.e., the home country). Formally, *consumer affinity* is defined as "*a feeling of liking, sympathy, and even attachment toward a specific foreign country that has become an in-group as a result of the consumer's direct personal experience and/or normative exposure and that positively affects the consumer's decision-making associated with products and services originating from the affinity country*" (Oberecker, Riefler, and Diamantopoulos, 2008, p. 26).

Riefler and Diamantopoulos (2007) and Oberecker, Riefler, and Diamantopoulos (2008) independently established that consumers indeed have affinities for certain countries. Addressing the bases of affinities, seven categories of affinity drivers that capture direct (i.e., so-called micro-drivers) and/or indirect (i.e., macro-drivers) experiences with a country can be identified. Microdrivers include individual interactions with the country like *traveling* to, having *contact* with (e.g., through friends or relatives), or long-term *stays* in the focal country while, macrodrivers capture a country's *lifestyle*, *scenery*, *culture*, and *politics and economics*.

Regarding the consumption-related consequences of consumer affinity, the insights gained from Oberecker, Riefler, and Diamantopoulos (2008) do not reveal a consistent positive impact of affinity on product judgment; however, findings point to a heightened willingness of particular consumer segments to buy products from the affinity country. This heightened willingness is largely the consequence of a desire to consume products both as a means for remaining connected to the affinity country and as a means for lowering perceived risk. Such segments (*see* MARKET SEGMENTATION AND TARGETING) might be also willing to pay price premiums for products from the focal country.

Consumer affinity can be differentiated from the constructs of xenophilia, internationalism,[4] and consumer animosity (for a detailed review, see Oberecker, Riefler, and Diamantopoulos, 2008). Rather than addressing favorable ATTITUDES toward foreign countries *in general* (like xenophilia and internationalism do), consumer affinity specifically refers to a positive, *country-specific* affect. Moreover, the consumer animosity construct, which captures negative sentiments toward a specific foreign country is not the polar opposite of consumer affinity, because, according to Oberecker, Riefler, and Diamantopoulos (2008), the key drivers of consumer affinity – which are a country's lifestyle and scenery as well as personal experiences through stays abroad – are different from Klein, Ettenson, and Morris's (1998) sources of consumer animosity (i.e., military, political, and economical events). Additionally, Jaffe and Nebenzahl (2006) expect that consumers rarely experience feelings of affinity and animosity for a specific country at the same time. Consumer affinity can be further distinguished from the country image construct[5] (*see* "COUNTRY OF ORIGIN" AS BRAND ELEMENT). While consumer affinity is conceptualized as being purely *affective* in nature and relating to a *specific* country, the country image construct comprises positive, negative, and neutral images for a variety of countries, the latter images being largely based on beliefs (i.e., cognitions). Furthermore, product country image[6] and consumer affinity are distinct

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

constructs. The former construct specifically addresses *product-related* characteristics like the product quality, whereas consumer affinity only reflects positive feelings toward a specific *country*.

Policy makers and marketing managers are advised to consider the consumer affinity segment when designing their marketing strategies (*see* MARKETING STRATEGY). They can undertake efforts in inducing such positive feelings (e.g., by attracting foreign students, tourists, or professionals to their country, or by launching commercials that highlight emotional sentiments toward a country) and thus leverage the heightened willingness to consume products and pay price premiums for products from the affinity country. At the same time, the consumer affinity segment could be a potential target segment in ethnocentric markets, where consumer preferences are likely to be dominated by domestic products.

ENDNOTES

[1] A *cosmopolitan consumer* is defined as "an open-minded individual, whose consumption orientation transcends any particular culture, locality, or community and who appreciates diversity including trying products and services from a variety of countries" (Riefler and Diamantopoulos, 2008).

[2] *Xenophilia* is defined as "love of strangers and foreigners and an implicit disrespect for one's own sociological reference group" (Perlmutter, 1954, p. 293).

[3] Affect itself is commonly seen as an "umbrella term for a set of more specific mental processes including emotions, moods [. . .] thus, affect might be considered a general category for mental feeling processes" (Bagozzi, Gopinath, and Neyer, 1999, p. 185). While cognition is, in general, equated with beliefs, affect equates with feelings (Baloglu and Brinberg, 1997).

[4] Internationalism is described as a concern about other nations' welfare and empathy for people of other countries (Kosterman and Feshbach, 1989).

[5] The *country image construct* is defined as "the total of all descriptive, inferential and informational beliefs one has about a country" (Martin and Eroglu, 1993, p. 193).

[6] The *product country image construct* is defined as "consumers' perceptions about the attributes of products made in a certain country" (Netemeyer, Bearden, and Sharma, 2003, p. 388).

Bibliography

Bagozzi, R., Gopinath, M., and Neyer, P. (1999) The role of emotions in marketing. *Journal of the Academy of Marketing Science*, **27** (2), 184–206.

Baloglu, S. and Brinberg, D. (1997) Affective images of tourism destinations. *Journal of Travel Research*, **35** (4), 11–15.

Jaffe E.D. and Nebenzahl I.D. (2006) It's all in the eyes of the consumer, in *National Image and Competitive Advantage: The Theory and Practice of Place Branding*, vol. **2** (eds E.D. Jaffe and I.D. Nebenzahl), Narayana Press, Copenhagen, pp. 79–109.

Klein, J.G., Ettenson, R., and Morris, M.D. (1998) The animosity model of foreign product purchase: an empirical test in the people's republic of China. *Journal of Marketing*, **62**, 89–100.

Kosterman, R. and Feshbach, S. (1989) Towards a measure of patriotic and nationalistic attitudes. *Political Psychology*, **10**, 257–274.

Martin, I. and Eroglu, S. (1993) Measuring a multi-dimensional construct: country image. *Journal of Business Research*, **28**, 191–210.

Netemeyer, R., Bearden, W.O., and Sharma, S. (2003) *Scaling Procedures*, Sage Publications, Thousand Oaks.

Oberecker, E., Riefler, P., and Diamantopoulos, A. (2008) The consumer affinity construct: conceptualization, qualitative investigation, and research agenda. *Journal of International Marketing*, **16** (3), 23–56.

Perlmutter, H. (1954) Some characteristics of the xenophilic personality. *The Journal of Psychology*, **38**, 291–300.

Riefler, P. and Diamantopoulos, A. (2007) Consumer animosity: a literature review and a reconsideration of its measurement. *International Marketing Review*, **24** (1), 87–119.

Riefler, P. and Diamantopoulos, A. (2008) Consumer cosmopolitanism: review and replication of the CYMYC scale. *Journal of Business Research*, **62** (4), 407–419.

Tajfel, H. (1982) Social psychology of intergroup relations. *Annual Reviews in Psychology*, **33**, 1–39.

global marketing ethics

Bodo B. Schlegelmilch

Global marketing ethics can be defined as *standards of conduct and moral judgment applied to the field of global marketing* (*see* GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES). As such, it is tightly interwoven with business ethics, in general, and marketing ethics, in particular. Notwithstanding the fact that business ethics has long been a central issue for many famous economists like Adam Smith, Max Weber, Frank Knight, and Friedrich August von Hayek, *marketing* ethics has played only a minor role in business schools and marketing journals until the 1960s (*see* MARKETING'S CORPORATE RESPONSIBILITY). The topics addressed during this period were primarily of a general nature, dealing with corporate ethical decision making or norms and codes, although some subfunctions of marketing were also debated at this early stage, including ethics in product management (*see* PRODUCT-LINE STRATEGIES; GLOBAL PRODUCT DEVELOPMENT), ethical issues in marketing research (*see* ETHICS IN MARKETING RESEARCH), or ethics in advertising (*see* INTERNATIONAL ADVERTISING – IS THERE STILL A STANDARDIZATION VERSUS LOCAL ADAPTATION DEBATE?).

However, it took until the 1980s for *global marketing ethics* to attract wider attention. Given that the international aspects of marketing ethics cause some of the most vexing dilemmas, this is remarkable. The early articles focusing specifically on international aspects of ethics were mainly comparative studies (*see* UNDERSTANDING COMPARATIVE MARKETING SYSTEMS THROUGH CHANNEL MAPPING). For example, Becker and Fritzsche (1987) contrasted ethical behavior of American, French, and German managers while Schlegelmilch (1989) discussed the ethics gap between the United Kingdom and United States. A literature review covering most of these contributions has been published by Tsalikis and Fritzsche (1989). Among the early books focusing on international aspects of business ethics is Donaldson's (1989) *The Ethics of International Business*. The increased focus on ethical issues in an international context also coincided with the launch of two specialized journals in the early 1980s, namely, the *Journal of Business Ethics* and the *Business* & *Professional Ethics Journal*.

The 1990s then witnessed a major expansion of the field. Contributions focused on a wide variety of international marketing issues like global marketing communication practices, ethical issues in international marketing research (*see* ETHICS IN MARKETING RESEARCH), price discrimination between international markets (*see* INTERNATIONAL PRICING OBJECTIVES AND STRATEGIES), counterfeits, use of corporate codes in an international marketing environment, and bribery. The first book exclusively devoted to *international* marketing ethics, was also published in the 1990s (Schlegelmilch, 1998).

Since 2000, the number of contributions on international marketing ethics increased substantially and the topics covered became more diverse. Issues as varied as ethical sensitivity to stakeholder interests across cultures (*see* MARKETING ASPECTS OF CULTURAL DISTANCE), international marketing ethics from an Islamic perspective, or the role of moral intensity and personal moral philosophies in the decision making of international marketers were discussed. While the large majority of previous contributions looked at ethical issues from a corporate vantage point, research now increasingly focuses on international marketing aspects through the lenses of consumers (*see* OPPORTUNITIES AND CHALLENGES IN SOCIAL MARKETING; ENVIRONMENTAL CONSUMER BEHAVIOR). Examples include comparisons of consumer ethics in different countries, attempts to draw up a typology of international consumers' ethical beliefs, or cross-cultural studies of the role of religion in consumers' ethical positions. A review of research focusing on marketing ethics was published by Nill and Schibrowsky (2007).

Taken collectively, international marketing ethics is still a relatively young and fragmented field. Given the growing interconnectedness of international markets, the increasing demands for corporate social responsibility, and the

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

ongoing calls for sound ethical leadership, more research in this area is needed.

Bibliography

Becker, H. and Fritzsche, D.J. (1987) A comparison of the ethical behavior of American, French and German managers *Columbia Journal of World Business*, (Winter), **22**, 87–95.

Donaldson, T. (1989) *The Ethics of International Business*, Oxford University Press, New York.

Nill, A. and Schibrowsky, J.A. (2007) Research on marketing ethics: a systematic review of the literature. *Journal of Macromarketing*, **27** (3), 256–273.

Schlegelmilch, B.B. (1989) The ethics gap between Britain and the United States: a comparison of the state of business ethics in both countries. *European Management Journal*, **7** (1), 57–64.

Schlegelmilch, B.B. (1998) *Marketing Ethics: An International Perspective*, International Thomson Publishing, London.

Tsalikis, J. and Fritzsche, D.J. (1989) Business ethics: a literature review with a focus on marketing ethics. *Journal of Business Ethics*, **8** (9), 695–743.

consumer world-mindedness

Susan P. Douglas and Edwin J. Nijssen

Definition

Consumer world-mindedness concerns an individual's interest in, openness to, and adoption of consumer products, services, and ideas from other cultures or parts of the world and accepting these cultures' norms and values completely and without *nationalistic bias or prejudice*. This follows Hannerz's (1990) view of a world-minded individual as one who has a predisposition toward a willingness to engage with the "other," and hence openness toward divergent cultural experiences, and a search for contrasts rather than for uniformity.

A world-minded person is someone with "positive respect for the culture and attainments of other races and peoples, an eagerness to know more about them, and a willingness to cooperate with them" (Fisher, 1926). World-minded individuals and consumers typically are more *interested in international media* and foreign travel and have a more developed international social network than their locally oriented counterparts (Beckmann *et al*., 2001). In addition, world-minded consumers are willing to *try out and experiment* with products and brands from other cultures to get to know and experience these cultures. They have a genuine openness and interest in other ideas and products, and as a result, have more positive attitudes toward foreign products and are more likely to purchase them.

Theoretical Roots and Related Concepts

"World-mindedness" has its origins in Socrates' concept of a "world citizen" and was first conceptualized empirically in the social science literature by Sampson and Smith (1957). They distinguish between international-mindedness and world-mindedness. *International-mindedness* is viewed as relating to interest in and knowledge about international affairs. The concept of world-mindedness in contrast, designates an individual who favors a worldview on problems, and whose primary reference group is mankind rather than Americans, English, Chinese, and so on. A world-minded individual is, therefore, likely to be concerned with human and animal rights, as also with poverty in emerging countries, world environment, and ecology.

Another related concept is that of *cosmopolitanism*. The concept of cosmopolitanism has been used in sociology. According to Merton (1957), a cosmopolitan is one who maintains a broad network of links and personal contacts outside the immediate community, while a localite is one whose universe and interests center on the local community. Hannerz (1990) further developed this concept, and conceptualized cosmopolitanism as a state of readiness or personal ability to make one's way into other cultures, accepting them completely and without prejudice. An in-depth study by Thompson and Tambyah (1999) views cosmopolitanism as the experience of an expatriate adapting to the lifestyle and mores of another country/culture without wholly abandoning his/her own. However, this latter view does not incorporate the adaptability of an individual to adapt to the world and its cultures in a broader sense, which as is the essence of world-mindedness.

Operationalization

Empirical studies in marketing typically rely on scales developed in other disciplines to measure world-mindedness or cosmopolitanism, but are generally not adapted to the consumer context. Sampson and Smith's (1957) 32 item scale of world-mindedness (or parts thereof) – relating to eight concepts, that is, religion, immigration, government, economics, patriotism, race, education, and war – is most often used for measuring consumers' world-mindedness. An exception is an instrument developed by Cannon *et al*., (1994) to profile a cosmopolitan based on items relating to community and organizational commitment as well as their tendency to define themselves in a broader world context. The scale was refined in a multicountry study by Yoon, Cannon, and Yaprak (1996), but the dimensionality and reliability of the scale remains open to question given low factor scores and reliabilities.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Nijssen and Douglas (2008) developed a two-dimensional second-order scale for operationalizing consumer world-mindedness, conceptualizing consumers' cultural openness, and cultural adaptability to products and services from other cultures as *formative indicators* of the underlying phenomenon. Differing from previous attempts drawing on Sampson and Smith, these authors have defined and adapted the construct to the consumer level and context.

Bibliography

Beckmann, S., Botschen, G., Botschen, M. *et al*. (2001) The world-minded consumer: an emic exploration, in *Advances in Consumer Research*, vol. **28** (eds M.C. Gilly and J. Meyers-Levy), Association for Consumer Research, Valdosta, 138.

Cannon, H.M. and Yaprak, A. (2002) Will the real world citizen please stand up! The many faces of cosmopolitan consumer behavior. *Journal of International Marketing*, **10** (4), 30–52.

Cannon, H.M., Yoon, S.-J., McGowan, L., and Yaprak, A. (1994) In search of the global consumer. Paper presented at the 1994 Annual Conference of the Academy of International Business.

Fisher, G.M. (1926) Does the Christian movement promote world-mindedness abroad? *Religious Education*, **21**, 179–187.

Hannerz, U. (1990), Cosmopolitans and locals in world culture, in *Theory, Culture and Society* (ed. M. Featherstone), Sage Publications, London, 237–251.

Merton, R.K. (1957) *Social Theory and Social Structure*, The Free Press, New York.

Nijssen, E.J. and Douglas, S.P. (2008) Consumer World-mindedness and social-mindedness and store image. *Journal of International Marketing*, **16** (3), 84–107.

Sampson, D.F. and Smith, H.P. (1957) A scale to measure world-minded attitudes. *The Journal of Social Psychology*, **45** (1), 99–106.

Thompson, C.J. and Tambyah, S.K. (1999) Trying to be cosmopolitan. *Journal of Consumer Research*, **26**, 214–241.

Yoon, S.J., Cannon, H.M., and Yaprak, A. (1996) A cross-cultural study of the cosmopolitan construct: validation of a new cosmopolitanism scale. Paper presented at the 1996 Annual Conference of the Academy of International Business.

global product R&D

Aysegul Özsomer

Global Product R&D refers to a firm's attempt to concentrate and coordinate its product-related R&D activities in a few countries (Porter, 1986; Zou and Özsomer, 1999). This means that R&D activities are undertaken with a focus on multiple customer groups from multiple types of markets (e.g., advanced and emerging). The development of products (and services) for the simultaneous consumption of multiple markets and multiple segments are strategic imperatives in a globalizing environment (*see* GLOBAL PRODUCT DEVELOPMENT).

Concentration enables a firm to benefit from the *location advantages* associated with different countries and coordination exploits cross-country synergies. *Coordination* captures the process of integrating product-related R&D across different subsidiaries around the world. A high level of coordination means that R&D activities are linked tightly across countries. Concentration refers to the conduct of product R&D activities only in a handful of counties (Zou and Cavusgil, 1996). A firm can choose to concentrate product R&D activities or it can choose to disperse them so that they are being replicated in many countries. The dispersed activities can still be highly coordinated via formal (e.g., IT intranets, frequent reporting) and *informal coordination* mechanisms (e.g., brainstorming sessions, frequent visits, multicountry teams).

There are two critical strategic concerns in implementing a global product R&D: (i) in which countries should the R&D activities be located so that location advantages (*see* MARKET ENTRY AND EXPANSION) can be exploited? (ii) How should these R&D activities be linked and coordinated to benefit from cross-country *synergies* (Porter, 1986; Roth, 1992)? Proper location selection or configuration allows the firm to exploit the comparative technological advantages of multiple countries, so that maximum effectiveness and efficiency can be attained (*see* COMPETITIVE ADVANTAGE). Coordination, on the other hand, establishes concerted action among R&D activities and is critical in managing interdependence between R&D activities and R&D centers. Coordination also enhances cross-country learning and scope benefits (Roth, 1992).

The shift to global product R&D by many multinational companies (MNCs) has led to the emergence of global *R&D Networks*. In addition to absolute increases in cross-border product R&D investments since the 1980s, the roles played by overseas R&D centers has changed from primarily transferring R&D from the parent company to the host country and adapting the product to the host country (*see* STANDARDIZATION/ADAPTATION OF INTERNATIONAL MARKETING STRATEGY) to that of a center for the development of product innovations for the *local*, *regional*, and *global markets* (Birkinshaw, Hood and Jonsson, 1998). Some of these product R&D centers have become "*Centers of excellence*" for MNCs.

The primary drivers of global product R&D networks are *differentiation* and division of labor. Different centers work on different product projects or on different activities for the same project leading to increased *specialization*. Such differentiation makes integration and coordination even more critical. It is possible that the benefits of global product R&D are related to the type and development stage of a technology. In industries characterized by rapid, radical innovations (*see* RADICAL INNOVATION), as opposed to incremental innovations (*see* INNOVATION TYPOLOGIES) and high uncertainties, the flexibility and speed enabled via global product R&D networks are critical.

The function of global product R&D may be related to whether product R&D is more *exploratory*, capturing search, variation, risk taking, experimentation, play, flexibility, discovery, and radical innovation, or whether it is more *exploitative*, capturing refinement, choice, production, efficiency, implementation, execution, and incremental innovation (March, 1991). When the function is more an exploratory global product R&D, interorganizational collaborations such as *partnerships* and *strategic alliances* are commonly resorted to in order to offset the higher risks and distant returns of the projects.

A critical research stream in global product R&D is *location selection* (*see* MARKET ENTRY

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

AND EXPANSION). Recent evidence in biotechnology and nanotechnology reveals that as research moves from *basic* to *applied*, the technological diversity, scientific strength, and scientific excellence of the host country are significant in attracting global product R&D centers of MNCs (Fernández-Ribas and Shapira, 2009). By contrast, the influence of local market-driven factors such as market size (Vernon, 1966) is less clear.

Bibliography

Birkinshaw, J.M., Hood, N. and Jonsson, S. (1998) Building firm-specific advantages in multinational corporations: the role of subsidiary initiative. *Strategic Management Journal*, **19**, 221–242.

Fernández-Ribas, A. and Shapira, P. (2009) Technological diversity, scientific excellence and the location of inventive activities abroad: the case of nanotechnology. *The Journal of Technology Transfer*, **34**, 286–303.

March, J.G. (1991) Exploration and exploitation in organizational learning. *Organization Science*, **21**, 71–87.

Porter, M.E. (1986) Patterns of international competition. *California Management Review*, **28**, 9–40.

Roth, K. (1992) International configuration and coordination archetypes for medium-sized firms in global industries. *Journal of International Business Studies*, **23**, 533–549.

Vernon, R. (1966) International investment and international trade in the product cycle. *Quarterly Journal of Economics*, **80**, 190–207.

Zou, S. and Cavusgil, S.T. (1996) Global strategy: a review and an integrated conceptual framework. *European Journal of Marketing*, **30**, 52–69.

Zou, S. and Özsomer, A. (1999) Global product R&D and the firm's strategic position. *Journal of International Marketing*, **7**, 57–76.

product ethnicity

Jean-Claude Usunier

Some products are typically associated with a particular country of origin, others are associated with several countries (e.g., cars with Germany, Japan, USA), while most products are not associated with specific countries. Products may be linked to a country because of its location, climate, or natural resources, or because of traditional manufacturing know-how (*see* "COUNTRY OF ORIGIN" AS BRAND ELEMENT). Countries also become associated with products because they are known as *place of invention*, *transformation*, or *use* (e.g., tea for England), or home to a brand associated with this country through its language (Leclerc, Schmitt, and Dubé, 1994).

These associations, dubbed by Roth and Romeo (1992) as *product–country matches*, may change consumer perception. A country considered as a poor origin, in general, may be viewed more positively when the particular product is associated with this country. Product ethnicity refers to the stereotypic association of a generic product with a particular country of origin (*see* CONSUMER CATEGORIZATION). Consumers from different countries may not make the same product ⇒ country or country ⇒ product associations. Thus, one may refer to the French product ethnicity of wine for German consumers, or to the Australian product ethnicity of wine for British consumers. Ethnicity in "product ethnicity" refers to a country (or several countries) being perceived as a legitimate place for the design, manufacturing, or consumption of a generic product (*see* PRODUCT CATEGORY). Countries may be associated with particular products given as stimuli (i.e., *product ⇒ country associations*) or products may be associated with particular countries given as stimuli (i.e., *country ⇒ product associations*). Product ethnicity occurs when an association reaches a certain level both ways. *Global product ethnicity* (GE), is based on country ⇔ product matches that are shared cross nationally. GE reflects the extent to which product ⇔ country associations are: (i) bidirectional and strong (i.e., most consumers make the association both ways), (ii) quasi-exclusive (i.e., the product is significantly associated with a single or very few countries-of-origin), and (iii) cross national or global (i.e., a particular country ⇔ product association is shared worldwide). Usunier and Cestre (2007) indicate how to calculate GE from associations made by respondents between products and countries (and vice versa).

Product ⇒ country associations have been shown to be related to consumer product familiarity and product involvement (*see* CONSUMER INVOLVEMENT), while country ⇒ product associations are related to country familiarity (Usunier and Cestre, 2007). As people are more familiar with their own country than with others, consumers from a particular country associate their country more frequently than other countries with product stimuli. Likewise, local consumers tend to associate more products with their own country than with other stimulus countries because they are exposed to local manufactures and brands (*see* PERCEPTION OF BRAND EQUITY), and they are more knowledgeable about and familiar with domestic products. Self-centered country ⇔ product associations reflect a positive "domestic country bias," in favor of one's own country and domestic products (Balabanis and Diamantopoulos, 2004). Such self-centered association tendencies have been shown to be stronger in individualist countries than in collectivist countries (Usunier and Cestre, 2007). Finally, consumer willingness to buy a particular product is positively related to the degree of congruence between its country of origin and its product ethnicity. Consequently, global marketers (*see* GLOBAL MARKETING STRATEGY; GLOBAL MARKETING STRATEGY: PERSPECTIVES AND APPROACHES; GLOBAL BRANDING: THREE KEYS FOR GLOBAL BRAND SUCCESS) should always assess the degree of product ethnicity for the product category for which a specific marketing strategy is being developed.

Bibliography

Balabanis, G. and Diamantopoulos, A. (2004) Domestic country bias, country-of-origin effects, and consumer ethnocentrism: a multidimensional unfolding app-

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

roach. *Journal of the Academy of Marketing Science*, **32** (1), 80–95.

Leclerc, F., Schmitt, B.H., and Dubé, L. (1994) Foreign branding and its effects on product perceptions and attitudes. *Journal of Marketing Research*, **31** (2), 263–270.

Roth, M.S. and Romeo, J.B. (1992) Matching product category and country image perceptions: a framework for managing country-of-origin effects. *Journal of International Business Studies*, **23** (3), 477–497.

Usunier, J.-C. and Cestre, G. (2007) Product ethnicity: revisiting the match between products and countries. *Journal of International Marketing*, **15** (3), 32–72.

born global

Gary Knight

"Born globals" are companies that conduct international business at or near the founding of the firm (Knight and Cavusgil, 1996, 2004; McKinsey and Company, 1993; Rennie, 1993). Despite the limited resources that usually characterize new businesses, born globals achieve substantial international sales from an early stage in their development. The period from domestic establishment to initial foreign market entry is often three or fewer years. Since the late 1980s, born globals have emerged in substantial numbers worldwide and have become relatively common among internationally active firms (*see* MARKET ENTRY AND EXPANSION). Historically, international business was mainly undertaken by large, well-resourced multinational enterprises (MNEs). The widespread appearance of born globals has revolutionized the traditional character of international business and holds important implications for the global economy.

Born globals are also known as *international new ventures*, which are defined in the scholarly management literature as "business organizations that, from inception, seek to derive significant competitive advantage from the use of resources and the sale of outputs in multiple countries" (Oviatt and McDougall, 1994). Given their youth, born globals tend to be relatively small firms and thus, usually have far fewer financial, human, and tangible resources than large MNEs. As the *internationalization* of born globals occurs shortly after founding, the origins of such firms are strongly international. Managers in these firms commit significant organizational resources to international activities. Born globals tend to be formed by entrepreneurs with a significant global orientation and often with strong international marketing skills (*see* GLOBAL MARKETING STRATEGY).

Born globals have been observed and described in virtually all major trading countries (OECD, 1997), across industry sectors (Knight and Cavusgil, 2004), and in both high- and low-tech industries (Madsen and Servais, 1997; Rennie, 1993). Historically, in many countries, it was believed that firms had to build a strong domestic base before venturing into foreign markets (*see* STRATEGIC EXPORT MARKETING–ACHIEVING SUCCESS IN A HARSH ENVIRONMENT). One reason was the high fixed costs of entering a new market at a distance, including the costs of gaining market information and of managing agents or representatives to establish effective sales organizations (*see* INTERNATIONAL MARKETING CHANNELS). However, dramatic changes in recent decades – primarily globalization and the emergence of key technologies in information, communication, and transportation – are facilitating the ability of companies to compete internationally in the earliest days after the founding of the firm (*see* DIGITAL MEDIUM AND GLOBAL MARKETING; FORCES AFFECTING GLOBAL INTEGRATION AND GLOBAL MARKETING).

Born globals are associated with the emergence of a new scholarly field, INTERNATIONAL ENTREPRENEURSHIP, which has been defined as a combination of innovative, proactive, and risk-seeking behaviors that crosses national borders and aims to create value in organizations (McDougall and Oviatt, 2000). The field emphasizes two principal streams of scholarly research: the growing international role played by young entrepreneurial firms (born globals) and the international, entrepreneurial activities of established firms (Zahra and George, 2002). In this way, research on born global firms is leading to findings on organizational behaviors and approaches that firms of any age can employ to undertake entrepreneurial international ventures.

Bibliography

Knight, G. and Cavusgil, S.T. (1996) The born global firm: a challenge to traditional internationalization theory, in *Advances in International Marketing* (eds S. Cavusgil and T. Madsen), vol. 8, JAI Press, Greenwich, pp. 127–137.

Knight, G. and Cavusgil, S.T. (2004) Innovation, organizational capabilities, and the born global firm. *Journal of International Business Studies*, **35** (2), 124–141.

Madsen, T. and Servais, P. (1997) The internationalization of born globals – an evolutionary process. *International Business Review*, **6** (6), 1–14.

McDougall, P. and Oviatt, B. (2000) International entrepreneurship: the intersection of two research

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

paths. *Academy of Management Journal*, **43** (5), 902–906.

McKinsey and Company (1993) Emerging Exporters: Australia's High Value-Added Manufacturing Exporters. Australian Manufacturing Council, Melbourne.

OECD (1997) Globalization and Small and Medium Enterprises (SMEs), Organization for Economic Co-operation and Development, Paris.

Oviatt, B. and McDougall, P. (1994) Toward a theory of international new ventures. *Journal of International Business Studies*, **25** (1), 45–64.

Rennie, M. (1993) Born global. *McKinsey Quarterly*, **4**, 45–52.

Zahra, S. and George, G. (2002) International entrepreneurship: the current status of the field and future research agenda, in *Strategic Leadership: Creating a New Mindset* (eds M. Hitt, R. Ireland, M. Camp, and D. Sexton), Blackwell, London, pp. 255–288.

export assistance programs

Esra F. Genctürk

Definition of Export Assistance Programs

The phrase *export assistance programs* refers to all public measures designed to encourage and assist exporting activites of individual firms and/or specific domestic industries. Almost all governments have a strong incentive to promote national exports in order to improve the international competitive advantage of domestic firms and hence the country's balance of trade. From a firm's perspective, such programs foster a pro-exporting attitude, facilitate development of necessary contacts, and provide assistance with specific exporting problems. Accordingly, variety of export assistance programs have been developed by governmental organizations to provide local firms/industries with support necessary for successful foreign market entry and expansion (*see* MARKET ENTRY AND EXPANSION) and to enhance their export performance (*see* EXPORT PERFORMANCE).

Types of Export Assistance Programs

Export assistance programs may be in many different forms ranging from counselling, tax incentives, and export financing to trade shows and sales leads. The kind of assistance that can be provided is restricted by the World Trade Organization (WTO) and as a result, government support typically manifests itself in a wide range of backup services (Diamantopoulos, Schlegelmilch and Tse, 1993). These services take different forms in different countries depending on a country's trade policy and its views toward government intervention in the business sector.

Apart from such differences, export assistance programs can be divided into indirect and direct services. The indirect services include all types of government support which are not specifically designed for exports but generate positive spillovers for export activity. These include productivity, R&D, and innovation support programs as well as general tax/investment incentive policies. The direct services, on the other hand, include all programs designed to enhance export activity and competitiveness. These direct services are, in turn, grouped into (i) financial acitivities, (ii) information services, and (iii) export facilitating activities (Albaum, Duerr and Strandskov, 2005; Hollensen, 2007).

Financial activities. Export assistance programs involving financial supports for exporting are direct government attempts to make expoting more attractive on economic grounds. They also include attempts to encourage greater participation in exporting, particularly by smaller firms. Granting of subsidies fall under this subgroup. While subsidies given only to exports would be a violation of WTO regulations, general subsidies may still enable domestic producers to become price-competitive overseas. However, provision of such subsidies not only results in threats of direct retaliation but may also lead to dumping investigations in foreign markets where subsidized exports adversely affect the local manufacturers. Apart from subsidies, tax benefits are also used to encourage exporting. While the nature and extent of tax benefits given to exporters varies from country to country, they include deferred taxes on export earning, long-term tax holidays for export profits, and low export-profit tax rates. National governments also encourage exporting through their credit policy and credit insurance programs. These programs include credit facilities to cover working capital needs of exporters as well as subsidies and guarantees for export credits offered through commerical banks. Credit insurance and guarantees also cover certain commerical and political risks associated with a given export transaction.

Information services. National governments are a critical source of much of the basic information upon which exporting decisions are based. Examples of government information services, based on primary or secondary data relevant for exporters, include country data, reports on foreign firms, OPPORTUNITY IDENTIFICATION bulletins, information on relevant home and host-government regulations, and identification of potential foreign buyers/distributors for various products in different countries.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Export facilitating activities. National governments also organize and sponsor activities in order to assist firms in tapping opportunities in foreign markets and expanding their foreign market coverage. These export facilitating activities include operating trade development offices or trade centers abroad, sponsoring trade missions, organizing trade fairs and exhibitions, and establishing free trade zones, free ports and/or export processing zones.

Bibliography

Albaum, G., Duerr, E. and Strandskov, J. (2005) *International Marketing and Export Management*, 5th edn, Prentice Hall, New York.

Diamantopoulos, A., Schlegelmilch B. and Tse, K. (1993) Understanding the role of export marketing assistance: empirical evidence and research needs. *European Journal of Marketing*, **27** (4), 5–18.

Hollensen, S. (2007) *Global Marketing*, 4th edn, Financial Times/Pearson Education Ltd., Harlow, UK.

international relationship marketing

Jaqueline Pels

Introduction

"Relationship Marketing: Shift or Shaft" was the provocative title of Jag Sheth's welcome speech during the Relationship Marketing Research Conference, at Emory University, back in 1996. Four years later, in the Handbook of Relationship Marketing (Sheth and Pravatiyar, 2000), in their last chapter, the authors addressed this challenging question repositioning it as "will it become a discipline out of a domain?" (p. 610). They avoided giving a clear answer and highlighted, instead, a set of challenges that relationship marketing needed to overcome, such as delimiting the domain, agreeing on a definition, developing a set of performance metrics, among others.

Almost 15 years later, the question still does not have an answer. Interestingly, a recent content analysis of the major business marketing journals (Dant and Lapuka, 2008; LaPlaca, 2008) shows that the share of articles dealing with relationships has increased consistently since 2000, whereas the relative number of articles on buying behavior, selling and sales management, and segmentation has been declining since the early 1990s. By studying the scholarly debates and the papers that have been published on this topic, it could be argued that there seems to be an academic shift toward relationship marketing. The question that needs to be addressed is this: "Is this shift occurring at a managerial level and, more specifically, is it an international trend?"

What does the International Empirical Data Say?

The Contemporary Marketing Practice (CMP) group has developed a strong conceptual classification scheme (Coviello *et al*., 1997; Pels *et al*., 2000; Brodie *et al*., 2006), as well as, a consistent body of empirical studies[1] involving several years (1997–2007) and multiple countries from Europe (e.g., UK, Germany, Finland, Sweden, Ireland, Spain, Russia), North America (e.g., Canada, USA), South America (e.g., Chile, Argentina, Uruguay), Asia (e.g., China, Thailand, New Zealand), and Africa (e.g., Ghana, Ivory Coast).

Brodie *et al*. (2008), in their history of the CMP project, describe how it was born as a result of a group of New Zealand researchers' perceived tension between theory and practice. Thus the program's objective was to *profile marketing practice in a contemporary environment, and to examine the relevance of relational marketing in different organizational, economic, and cultural contexts*. The final aim was to develop an understanding of how firms relate to their markets in a manner that integrates both traditional and more modern views of marketing. A classification scheme was developed that identified a typology of four[2] marketing practices: transaction marketing (*see* MARKETING MIX), database marketing (*see* INTERNAL DATABASES FOR MARKETING RESEARCH; DATABASE MINING AND MARKETING), interaction marketing (*see* INTERNATIONAL BUSINESS-TO-BUSINESS MARKETING), and network marketing.

In their article, Coviello et al. (2002) shift the analysis from "fitting data into the typology" to "allowing the data to speak." The authors applied a cluster analysis and suggested a taxonomy of three marketing practices: a transactional cluster, a relational cluster, and a transactional/relational cluster (see Table 1). When replicated for Latin America (Pels and Brodie, 2003; Pels *et al*., 2004), for Russia (Wagner, 2005), and for Ghana and Ivory Coast (Dadzie *et al*., 2008) (*see* EMERGING MARKETS) a fourth cluster was added to the initial CMP taxonomy – the low-marketing cluster (see Table 2).

Conclusion

Relationship marketing suggests a different way of understanding the environment (i.e., as networks rather than markets) and, as a result, a diverse way of interacting with it (i.e., through close, interactive exchanges in which value is cocreated by the parties involved). Both the academic as well as the managerial world have much to gain in studying and understanding these types of exchanges. Nevertheless, the empirical studies do not support a shift toward

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Table 1 CMP cluster and types of marketing practice (Canada, USA, New Zealand, Finland, and Sweden).

Cluster	*Transactional (n = 103)*	*Transactional /Relational (n = 107)*	*Relational (n = 98)*	*Mean Score (All 308 Firms)*
Transaction Marketing Score	*0.81*	*0.85*	0.65	0.79
Database Marketing Score	0.63	*0.78*	0.60	0.68
Interaction Marketing Score	0.63	*0.82*	*0.79*	0.75
Network Marketing Score	0.48	*0.75*	*0.71*	0.64

Source: Adapted from Coviello, N. E., Brodie, R. J., Danaher, P. J. and Johnston, W. J. (2002) How firms relate to their markets: An empirical examination of contemporary marketing practice. Journal of Marketing, Vol. 66, No. 3, 33–46.
Note: Scores range between 0 and 1. Scores above 0.80 reflect higher levels of marketing practice; scores between 0.61 and 0.80 reflect medium levels of marketing practice; scores less than 0.61 reflect lower levels of marketing practice.
Note: numbers in italics represent higher than mean scores.

Table 2 CMP cluster and types of marketing practice (Argentina).

Clusters	*Pluralistic*	*Relationship and Network*	*Low Marketing*	*Mean Score (All Firms)*
Transaction	*0.68*	0.51	0.60	0.61
Database	*0.67*	0.46	0.51	0.57
Interaction	*0.80*	*0.82*	0.61	0.76
Network	*0.73*	*0.71*	0.51	0.68
% of firms	32%	50%	18%	—

Source: Pels, J., Brodie, R. J. and Johnston, W. J. (2004) Benchmarking business-to-Reproduced from business practices in emerging and developed economies: Argentina compared to the USA and New Zealand. Journal of Business and Industrial Marketing, Vol. 19, No. 6, 386–396.

relationship marketing and we need to be very care cautious before proclaiming the universality of this view. In this postmodern world where diversity is the rule we, academics, would serve the business community better if we start by acknowledging the diversity in marketing modes!

Endnotes

[1] See the CMP web page for complete list of empirical articles http://cmp.auckland.ac.nz/.

[2] A fifth e-marketing practice was later added, see Coviello et al. (2001).

Bibliography

Brodie, R.J., Coviello, N.E., and Winklhofer, H. (2008) Contemporary marketing practices research program: a review of the first decade. *Journal of Business and Industrial Marketing*, **23** (2) (Special Issue), 84–94.

Brodie, R.J., Pels, J., and Saren, M. (2006)From goods-towards service-centered marketing: dangerous dichotomy or an emerging dominant logic? in *The New Dominant Logic of Marketing* (R., Vargo and S., Lusch), M.E. Sharpe, pp. 307–319.

Coviello, N.E., Brodie, R.J., Danaher, P.J., and Johnston, W.J. (2002) How firms relate to their markets: an empirical examination of contemporary marketing practice. *Journal of Marketing*, **66** (3), 33–46.

Coviello, N.E., Brodie, R.J., and Munro, H.J. (1997) Understanding contemporary marketing: development of a classification scheme. *Journal of Marketing Management*, **13** (6), 501–522.

Coviello, N.E., Milley, R., and Marcolin, B. (2001) Understanding IT-enabled interactivity in contemporary marketing. *Journal of Interactive Marketing*, **15** (4), 18–33.

Dadzie, K., Johnston, W.J., and Pels, J. (2008) Business-to-business marketing practices in emerging economies: West Africa and Argentina bench marked with the United States. *Journal of Business In Industrial Markets*, **23** (2), 115–123.

Dant, R.P. and Lapuka, I.I. (2008) The Journal of Business-to-Business Marketing comes of age: some postscripts. *Journal of Business-to-Business Marketing*, **15** (2), 192–197.

LaPlaca, P.J. (2008) Commentary on The essence of business marketing . . . by Lichtental, Mummalaneni, and Wilson: the JBBM comes of age. *Journal of Business-to-Business Marketing*, **15** (2), 180–191.

Pels, J. and Brodie, R.J. (2003) Profiling marketing practice in an emerging economy: the Argentine case. *Journal of Global Marketing*, **17** (1), 67–91.

Pels, J., Brodie, R.J., and Johnston, W.J. (2004) Benchmarking business-to-business practices in emerging and developed economies: Argentina compared to the USA and New Zealand. *Journal of Business and Industrial Marketing*, **19** (6), 386–396.

Pels, J., Coviello, N.E., and Brodie, R.J. (2000) Integrating transactional and relational marketing exchange: a pluralistic perspective. *Journal of Marketing Theory and Practice*, **8** (3), 11–20.

Sheth, J.N. and Parvatiyar, A. (2000) *Handbook of Relationship Marketing*, Sage Publications, Thousand Oaks.

Wagner, R. (2005) Contemporary marketing practices in Russia. *European Journal of Marketing*, **39** (1/2), 199–215.

international entrepreneurship

Nicole Coviello

Research in the area of international entrepreneurship (IE) stems in large part from McDougall's (1989) empirical study comparing domestic and international new ventures (INVs), and Oviatt and McDougall's (1994) seminal work, which provided a theoretical base for the study of INVs. *INVs* were defined as business organizations "that, from inception, [seek] to derive significant competitive advantage from the use of resources and the sale of outputs in multiple countries" (Oviatt and McDougall, 1994, p. 49). Thus, study in the emerging field of IE began with an interest in new ventures.

The first widely accepted definition of IE comes from McDougall and Oviatt (2000, p.-903). They defined *IE* as " ... a combination of innovative, proactive and risk-seeking behavior that crosses national borders and is intended to create value in organizations." This was later revised to reflect theoretical advances in the parent discipline of entrepreneurship, particularly those pertaining to "opportunity." Consequently, *IE* was redefined as " ... the discovery, enactment, evaluation, and exploitation of opportunities – across national borders – to create future goods and services" (Oviatt and McDougall, 2005, p. 540). By allowing for individual, group, and organizational levels of behavior, the field of IE was broadened from its early focus on new ventures to include the behavior of larger, more established firms and the concept of corporate entrepreneurship. This is consistent with Zahra (2005), who maintains that it is the firm's entrepreneurial actions and entrepreneurial qualities that are the major source of competitive advantage. Also found in the domain of IE are comparisons of entrepreneurial behavior in multiple countries and cultures (Coviello and Jones, 2004; Oviatt and McDougall, 2005).

Nevertheless, central to IE research is the INV. Successful INVs usually lack resources and, consequently, internalize few assets and focus on less costly governance mechanisms to control a greater percentage of vital resources than would mature organizations. Such governance mechanisms include network structures and social control within cooperative relationships. Further, successful INVs gain foreign location advantages from private knowledge they possess or produce, and make it sustainable through one or more means of protection. These include imperfect imitability, licensing, networks, or patents (Oviatt and McDougall, 1994). As discussed by Zahra (2005), the INV assumptions and cognitions regarding markets and competition are different from those of well-established firms. INV theory also challenges the more risk-averse, constrained, and gradual patterns of internationalization posited by Johanson and Vahlne's (1977) process theory (Autio, 2005).

Also relevant is that Oviatt and McDougall (1994) conceptualized four types of INV: (i) export/import start-ups, (ii) multinational traders, (iii) geographically focused start-ups, and (iv) global start-ups. It is the latter form that has received the most attention, commonly referred to as the BORN GLOBAL (Knight and Cavusgil, 1996). Zahra (2005) suggests, however, that insufficient IE research has attended to the four different types of INV, their prevalence under different combinations of industry, market, firm, and entrepreneur-related conditions or changes (including learning) that occur in INVs over time.

Recent theoretical arguments provide a base for further research in IE. Examples include Dimitratos and Plakoyianniki (2003)'s work on the dimensions of international entrepreneurial culture, Jones and Coviello's (2005) development of a general model of entrepreneurial international behavior and Styles and Seymour's (2006) arguments that value is not only created but exchanged in the process of identifying and exploiting opportunities that cross national borders. Such developments reflect the richness of IE as a research area, and its position at the interface of many disciplines ranging from international business, entrepreneurship, and strategy to, for example, marketing, innovation, knowledge management, and learning theory.

Bibliography

Autio, E. (2005) Creative tension: the significance of Ben Oviatt's and Patricia McDougall's article 'Toward

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

a theory of international new ventures'. *Journal of International Business Studies*, **36**, 9–19.

Coviello, N.E. and Jones, M.V. (2004) Methodological issues in international entrepreneurship research. *Journal of Business Venturing*, **19**, 485–508.

Dimitratos, P. and Plakoyianniki, E. (2003) Theoretical foundations of an international entrepreneurial culture. *Journal of International Entrepreneurship*, - **1**, 187–215.

Johanson, J. and Vahlne, J.-E. (1977) The internationalization process of the firm: a model of knowledge development and increasing foreign market commitment. *Journal of International Business Studies*, - **8**, 23–32.

Jones, M.V. and Coviello, N.E. (2005) Internationalization: conceptualizing an entrepreneurial process of behavior in time. *Journal of International Business Studies*, **36**, 284–303.

Knight G.A. and Cavusgil S.T. (1996) The born global firm: a challenge to traditional internationalization theory, in *Advances in International Marketing*, - vol. **8** (eds S.T. Cavusgil and T. Madsen), JAI Press, Greenwich, pp. 11–26.

McDougall, P.P. (1989) International versus domestic entrepreneurship: new venture strategic behavior and industry structure. *Journal of Business Venturing*, - **4**, 387–399.

McDougall, P.P. and Oviatt, B. (2000) International entrepreneurship: the intersection of two research paths. *Academy of Management Journal*, **43**, 902–908.

Oviatt, B.M. and McDougall, P.P. (1994) Toward a theory of international new ventures. *Journal of International Business Studies*, **25**, 45–64.

Oviatt, B.M. and McDougall, P.P. (2005) Defining international entrepreneurship and modeling the speed of internationalization. *Entrepreneurship Theory and Practice*, **29**, 537–553.

Styles, C. and Seymour, R.G. (2006) Opportunities for marketing researchers in international entrepreneurship. *International Marketing Review*, **23**, 126–145.

Zahra, S.A. (2005) A theory of international new ventures: a decade of research. *Journal of International Business Studies*, **36**, 20–28.

international business-to-business marketing

Paul Matthyssens

International business-to-business (B2B) marketing can be conceived as "cross-border value creation and exchange processes in business markets." These days, most B2B companies are engaged in some form or other of international marketing and selling – the so-called "outward" internationalization (Hollensen, 2007). Several factors trigger this internationalization process. The first trigger is the *delocalization of their customers*. In fact, clients are gradually moving their production operations to offshore locations (*see* OFFSHORING AND MARKETING). Suppliers often get requests for "client following" to remote countries and their traditional "Western" markets dry up. Business markets have become international markets and international customers claim the same level of service in all foreign locations.

A second driver is the *growth in international trade* and the opening up and development of new foreign markets. Foreign markets are growing rapidly and the WTO is facilitating entry in these markets (for a discussion on barriers, *see* EMBARGOES AND SANCTIONS). At the same time, national governments, trade associations, and chambers of commerce are often "pushing" their homegrown champions to consider expansion abroad (Hollensen, 2007).

The third trigger is the *internationalization and global coordination of the purchasing function* (*see also* GLOBAL SOURCING STRATEGY: AN EVOLUTION). Increasingly, purchasers from multinational companies seek global purchasing synergies. This way, local plants gradually lose their purchasing mandates. Suppliers might be forced to sell twice and at different levels: an international frame contract to the client's central purchasing office abroad, and sales to the local plant in the supplier's domestic market. Getting access to international headquarters might be very difficult for smaller "local" suppliers who might lack the necessary scale and references to respond to a call for quotation at the central level. International customers often prefer to work with suppliers with streamlined service approaches, integrated supply chain management, and internationally standardized marketing policies and GLOBAL ACCOUNT MANAGEMENT: THE RATIONALE AND MOTIVATION.

The fourth trigger is the *growing importance of international networks* (Harris and Wheeler, 2005). Solution providers (Helander and Möller, 2007) and value innovators (Matthyssens *et al.*, 2006) in B2B markets often rely on international partners for complementary competences. However, seeking good global partners and managing cross-border alliances turns out to be very challenging.

The fifth driver for increased internationalization is the intensified use of *electronic forms of exchange* (Samiee, 2008). Today, the Internet is a key facilitator for identifying suppliers or customers, negotiating, activating communities, establishing e-market places, and so on. Although entry and expansion in the international market are made much easier this way, B2B relationships might be more difficult to manage in the electronic highway (Samiee, 2008).

Because of these drivers, B2B companies recognize that international activities are fundamental to their performance (Katsikeas, 2006), and they start internationalizing earlier in their development (the so-called "BORN GLOBAL", Rialp *et al.*, 2005) and they expand internationally at a high speed. Insufficient resources lead B2B marketers to use intensely international networks.

The higher degree of internationalization puts pressure on B2B companies. They face the challenge of coordinating and streamlining their marketing processes and programs across countries and regions. They must seek global synergies in an effort to control costs and eliminate inconsistencies in their marketing approaches, while simultaneously being forced to be responsive to local habits (read further in *see* STANDARDIZATION/ADAPTATION OF INTERNATIONAL MARKETING STRATEGY). The result is a subtle marketing strategy approach of combined adaptation, aggregation, and arbitrage (Ghemawat, 2007).

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

Bibliography

Ghemawat, P. (2007) Managing differences: the critical challenge of global strategy. *Harvard Business Review*, **85** (3), 58–68.

Harris, S. and Wheeler, C. (2005) Entrepreneurs' relationships for internationalization: functions, origins and strategies. *International Business Review*, **14** (2), 187–207.

Helander, A. and Möller, K. (2007) System supplier's customer strategy. *Industrial Marketing Management*, **36** (6), 719–730.

Hollensen, S. (2007) *Global Marketing*, 4th edn, Prentice Hall, Harlow.

Katsikeas, C. (2006) Global marketing of industrial products: contemporary developments and future directions. *Industrial Marketing Management*, **35** (5), 540–544.

Matthyssens, P., Vandenbempt, K., and Berghman, L. (2006) Value innovation in business markets. Breaking the industry recipe. *Industrial Marketing Management*, **35**, 751–761.

Rialp, A., Rialp, J., and Knight, G. (2005) The phenomenon of early internationalizing firms: what do we know after a decade (1993–2003) of scientific inquiry? *International Business Review*, **14** (2), 147–166.

Samiee, S. (2008) Global marketing effectiveness via alliances and alectronic commerce in business-to-business markets. *Industrial Marketing Management*, **37** (1), 3–8.

smuggling

Kate Gillespie

Smuggling encompasses the import or export of goods in violation of domestic or international law. Most smuggling, however, involves the illegal import of goods that face import quotas, significant tariffs, or other high taxes (*see* EMBARGOES AND SANCTIONS).

During the mid-nineteenth century, most developing countries employed strict quotas or enacted high tariffs on imports to protect local industries. These policies limited product choice and increased prices on many products. As a result, smuggling of consumer products became widespread. Smugglers evaded quotas and avoided tariffs, thus providing consumers with more choices and cheaper prices. Smugglers co-opted law enforcement agents via bribery and utilized the many unlicensed street vendors of the developing world's informal economy for the retail distribution of smuggled products.

Trade liberalization in EMERGING MARKETS has decreased smuggling but has failed to eradicate it completely. Smugglers continue to avoid residual tariffs as well as other taxes that their legal counterparts are subject to such as income taxes, sales taxes, value-added taxes, and sin taxes. Some smugglers, in the face of trade liberalization, have decreased costs by shifting from the smuggling of legally purchased goods to the smuggling of stolen or counterfeit goods (Gillespie and McBride, 1996). Furthermore, as customs authorities in many countries improve their ability to identify and confiscate shipments of counterfeits, many counterfeiters are forced to employ smuggling channels to circumvent customs checks. Thus even in the third millennium, smuggling dominates trade between some countries (Golub and Mbaye, 2008).

Smuggling of common consumer products has come increasingly under the auspices of transnational organized crime. In the late 1990s, Colombian officials estimated that 45% of the imported consumer products sold in their country had been financed by the Black Market Peso Exchange, a nefarious institution in which smuggled goods help launder illicit drug money. Government attempts to eliminate smuggling have increased as a result of the consolidation and criminal violence of smugglers (Gillespie, 2003). Nonetheless, smuggling remains hard to eradicate because of its organized nature, government corruption, and a lack of resources for law enforcement.

There are both benefits and disadvantages for producers of legitimate products that are smuggled. Smuggling works to their advantage when their products are barred from entering markets or when prohibitive taxation discourages purchase of those products. However, smuggling poses similar challenges to multinational corporations as do legal but unauthorized parallel imports or GRAY MARKETS. Owners of global brands (*see* GLOBAL BRANDING: THREE KEYS FOR GLOBAL BRAND SUCCESS) may suffer when smugglers undercut legitimate distributors, since legitimate distributors often provide greater oversight of product quality and deliver after-sales service. Relationships with strategic allies, such as licensees or joint-venture partners in local markets, can be undermined when contraband enters partners' markets and diminishes the legal sales of these partners. However, the line between legal importers of products and smugglers is often blurred, because legal importers may distribute smuggled goods as well. Similarly, smugglers may attempt to camouflage their activities by legally importing some portion of the products they distribute.

Despite the challenges smuggling poses for international marketers, multinational firms have been known not only to tolerate the smuggling of their products but to actively support such smuggling. In the case of the world's five largest tobacco companies, support of smuggling has included selling cigarettes to known smugglers, disguising contraband shipments, destroying records related to smuggled cigarettes, establishing international financial accounts to camouflage contraband earnings, and purchasing airtime on transnational media to provide advertising support for smuggled brands. Several national and provincial governments have sued such companies for tax evasion and also to recoup costs incurred by the governments' policing of smuggling (Gillespie, 2003). Issues of jurisdiction have made such law suits difficult to win, but tobacco companies have sometimes settled out of court for sums rumored to be as high as a billion

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

dollars. Furthermore, as tobacco firms began to curtail supplies to distributors involved with smuggling, these distributors replaced the legitimate product with counterfeits. Thus smuggling ceased to be a major contributor to the profits of these companies and became, instead, a substantial threat to their sales and brands.

Bibliography

Gillespie, K. (2003) Smuggling and the global firm. *Journal of International Management*, **9**, 317–333.

Gillespie, K. and McBride, J.B. (1996) Smuggling in emerging markets: global implications. *Columbia Journal of World Business*, **31** (4), 39–54.

Golub, S.S. and Mbaye, A.A. (2008) National trade policies and smuggling in Africa: the case of the Gambia and Senegal. *World Development*, **37**, 595–606.

understanding comparative marketing systems through channel mapping

Matthew B. Myers and Joachim (Joe) Grass

Introduction

Despite a general increase in economic wealth worldwide, significant differences remain in levels of economic development throughout the globe. A disequilibrium of resource distribution is particularly evident in many transitional and developing economies (TDEs), which causes significant portions of the expanding population in many of these economies to "gradually shift into urban slums and create a proletariat" (Slater and Jenkins, 1979). Given the assumption that economic stability leads to social stability, and economic stability is only possible when stakeholders in a society have access to markets, understanding how to better distribute wealth in TDEs is a key factor in comparative marketing system research. Having said this, researchers have also failed to provide meaningful evidence relative to what type, or combination, of resource distribution is best suited to bring stability to markets and economies. Often, redistribution has been accomplished through either high tax levies, or worse, nationalization of properties. Instead, new perspectives within global marketing seek to employ the entrepreneur as the redistributive agent, specifically by supporting and utilizing the distribution function to facilitate more efficient and effective markets. Marketing, and especially distribution, is not only a function within individual firms or institutions, but rather part of the whole social process system of a society (Slater and Jenkins, 1979). According to the activist paradigm, which states that marketing is a stimulus of development, this means that marketing institutions can facilitate economic development (Taylor and Omura, 1994), and hence help remedy this condition.

Various factors are important for business development in TDE countries and the subsequent competitive positions these markets may enjoy relative to more developed economies. First, price controls must be addressed. Second, levels of technological development must be achieved. And third, infrastructure and distribution channel development must reach a point where the free flow of goods and services, and the corresponding redistribution of wealth internal to the market, takes place. It is the channel variable which spans a broad spectrum of issues ranging from availability of credit to natural resources (Albaum and Peterson, 1984). A useful yet underutilized tool for analyzing channels is "channel mapping" (Nason, 1993). Channel mapping identifies anomalies in a market's distribution processes and is used to increase efficiencies in distribution systems as well as propose market reforms that will factor in cultural variables to minimize detrimental effects on the market. This is particularly important in transitional countries where command systems have customarily ignored critical aspects of the distribution process (Nason, 1993). This can lead to inefficient, inflexible, and poorly structured channels, which can hinder economic development of an area (Slater *et al.*, 1969), restrict individual participation in the economy, and restrict foreign direct investments (FDI). Most importantly, success of reforms is in part dependent on the inclusion of individuals (entrepreneurs) in improving distribution efficiency, this by performing warehousing, retail, logistics, or other functions key with moving goods to and from markets.

Understanding Market Differentials through Channel Mapping

Analyzing comparative marketing systems in order to identify distribution anomalies may take place through a systematic process developed by Slater and described by Nason (1993). There are three steps required, starting with the analysis of a system and leading to market and policy reforms (see Figure 1).

Step 1: system analysis. The process starts with determining the flow of goods, services, and information through the various distribution channels, which helps identify and remedy inefficiencies in the system (such as insufficient or overabundance of middlemen, insufficient/deficient roads, raw materials, production and manufacturing, transportation, wholesale, warehousing, power relationships, consumers,

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

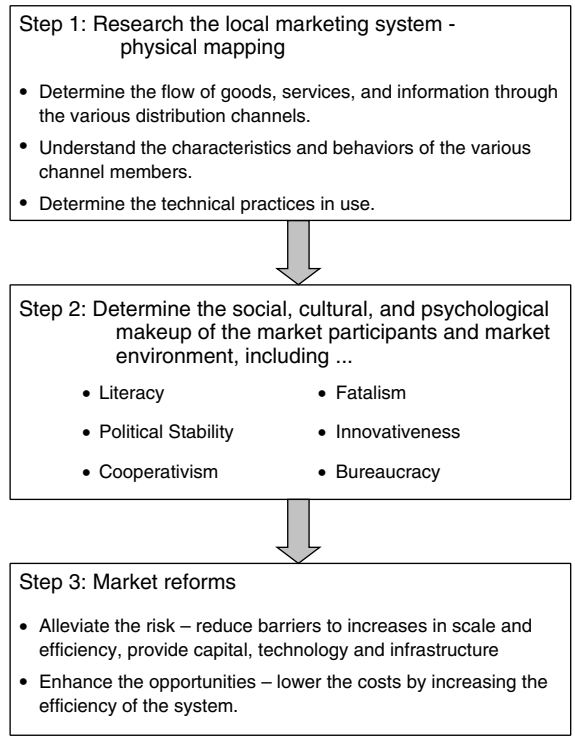


Figure 1 The channel mapping process.

and credit). This causes prices to decline and purchasing power of consumers to increase. This step also calls for understanding the characteristics and behaviors of the various channel members, which is especially important in an international business context so that potential foreign investors can better understand the economic environment and thereby reduce their risk.

Step 2: social and cultural environment. The second step calls for an analysis of the social, cultural, and psychological make-up of the market participants and environment (such as literacy rates, achievement desires, cooperativism, fatalism, innovativeness, and gender roles). This is necessary to understand the expected business and social behaviors in a certain environment, as well as to understand possible repercussions across the society that result from market reforms and new policies. From both standpoints (sociological and economic) it is important to understand barriers and perceived barriers of the channel participants.

Step 3: market reforms. The final step involves the creation of new policies and regulations that will (i) enhance the opportunities of the individual within the market through lowering costs (via increasing efficiencies), (ii) include an increasingly broad cross-section of the population in the value-added chain, and enhance the risk assessment capabilities for (domestic as well as international) business development.

Applications of the channel mapping method have been successful in a number of TDEs including Lesotho, Brazil, and Thailand. It is an effective method for identifying limitations on channel performance, creating greater participation in marketing functions and thus redistributing wealth, and focusing

developmental efforts at the micro- and macrosystem level on the most cost efficient rationalization strategies (Nason, 1993). Given the importance of economic and social stabilization worldwide, an increased understanding of the role of marketing distribution in the process of economic development is timely. Utilizing system comparisons such as channel mapping offers policymakers a viable option for market-friendly wealth redistribution, with benefits from the resulting increased participation in the marketplace.

Bibliography

Albaum, G. and Peterson, R.A. (1984) Empirical research in international marketing. *Journal of International Business Studies*, **15**, 161–173.

Nason, R. (1993) *Channel Mapping: Analysis for Market Development in Transition Economies*. Eighteenth Annual Macromarketing Conference, The University of Rhode Island.

Slater, C.C. and Jenkins, D.G. (1979) *System Approaches to Comparative Macro-Marketing. Macro-Marketing: New Steps on the Learning Curve* (eds G. Fisk and R.W. Nason), Graduate School of Business Administration, University of Colorado, Boulder, pp. 371–380.

Slater, C.C., Riley, H., Farace, V. *et al.* (1969) *The Market Processes of Recife, Brazil, and Its Food Shed*, Research Report No. 2, Latin American Studies Center, Michigan State University, East Lansing.

Taylor, C.R. and Omura, G.S. (1994) An evaluation of alternative paradigms in marketing and economic development – part I. *Journal of Macromarketing*, **14** (2), 6–20.

offshoring and marketing

Neeraj Bharadwaj

Background

Offshoring, which refers to the migration of value-chain processes once carried out in a given locale to another country,[1] has been deemed a practice likely to bring about transformational change in international trade (Blinder, 2006). Such a claim does not seem too far-fetched because many firms have readily transferred activities such as production, new-product development (*see* GLOBAL PRODUCT DEVELOPMENT), and pre- and post-sales assistance (*see* SERVICES MARKETING STRATEGY) to locations abroad, mainly enticed by the promise that offshoring will enable them to carry out these activities in the very same way, but at a lower cost (Friedman, 2005).[2]

Although some research in operations management (Gans, Koole, and Mandelbaum, 2003; Carmel and Tjia, 2005) as well as by prominent consulting firms (Farrell and Rosenfeld, 2005) suggests that offshoring has indeed assisted firms in reducing costs, the reality is that cost savings alone are unable to capture the full implications of the practice. That is because customers may perceive the quality of offshored "products" – be they goods and/or services – as subpar (Whoriskey, 2008). This is problematic because prior research has shown that a firm whose product/service quality is perceived as subpar is likely to encounter unfavorable customer evaluations and higher churn rates (Keaveney, 1995), which in turn, can have deleterious top-line revenue effects (Zeithaml, Berry, and Parasuraman, 1996).

It is therefore important to examine the research in marketing that has already investigated the impact of the geographic location cue on customer appraisals. It is also necessary to identify future projects that can be carried out to learn more about offshoring, customer appraisals, and cost savings.

Offshoring and Customer Appraisals

The country of origin (COO) literature (*see* "COUNTRY OF ORIGIN" AS BRAND ELEMENT) has studied the impact of the location cue on customer appraisals. In general, this body of knowledge demonstrates that favorable country perceptions yield favorable product/service evaluations, and advises that the location cue ought to be evaluated in conjunction with other cues (Peterson and Jolibert, 1995; Gurhan-Canli and Maheswaran, 2000). In one of the few studies to directly examine the influence of offshoring on customer appraisals, Roggeveen, Bharadwaj, and Hoyer (2007) build upon the COO literature to evaluate the effect of call-center location and firm reputation on customer expectations regarding an upcoming voice-to-voice service encounter. Those authors conclude that call-center location does not impact customer expectations regarding an upcoming service encounter if the firm is reputed; however, if the firm is less known, individuals anticipate being less satisfied if they believe the call center is located in a nation dissimilar to the United States (i.e., India or the Philippines) versus either similar to the United States (i.e., England) or in the United States.

Offshoring, Customer Appraisals, and Cost Savings

More scholarly attention should be devoted toward understanding the outcomes of this widespread practice of offshoring. One possibility is to further examine the impact of the location cue on customer evaluations, where location could be operationalized along the lines of cultural distance (*see* MARKETING ASPECTS OF CULTURAL DISTANCE) between the customer's home country and the country to which the business process is to be migrated. Bharadwaj and Roggeveen (2008) describe a methodology that can be utilized to evaluate situations in which the customer possesses greater familiarity with the culture to which a business is being migrated (i.e., "nearshoring") compared to a less familiar one. A more challenging endeavor is to integrate thinking in operations and marketing to evaluate the extent to which the short-term cost savings from offshoring are offset by top-line revenue losses stemming from existing customers exiting relationships, buying less, and/or engaging in negative word-of-mouth behavior and from prospective customers being

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

put off by perceived product/service quality inadequacies. Such research that evaluates the trade-off between initial cost savings and long-term customer-related outcomes will provide insight into the conditions under which offshoring can truly serve as a basis for driving profitable growth.

Endnotes

[1] Bharadwaj and Roggeveen (2008) detail the conceptual distinction between offshoring (i.e., when a given business process is carried out somewhere else) and outsourcing (i.e., when a given process is carried out by someone else). Researchers must distinguish between these two closely related practices in order to understand their respective main and interaction effects.

[2] Another reason firms opt to engage in offshoring is to draw upon new employee skill sets available elsewhere (Venkatraman, 2004).

Bibliography

Bharadwaj, N. and Roggeveen, A.L. (2008) The impact of offshored and outsourced call service centers on customer appraisals. *Marketing Letters*, **19**, 13–23.

Blinder, A.S. (2006) Offshoring: the next industrial revolution? *Foreign Affairs*, **85** (2), 113–128.

Carmel, E. and Tjia, P. (2005) *Offshoring Information Technology: Sourcing and Outsourcing to a Global Workforce*, Cambridge University Press, Cambridge.

Farrell, D. and Rosenfeld, J. (2005) U.S. offshoring: rethinking the response, in *Offshoring: Understanding the Emerging Global Labor Market* (ed. D. Farrell), Harvard Business School Press, Boston, pp. 123–142.

Friedman, T.L. (2005) *The World is Flat*, Farrar, Strauss and Giroux, New York.

Gans, N., Koole, G., and Mandelbaum, A. (2003) Telephone call centers: tutorial, review, and research prospects. *Manufacturing and Service Operations Management*, **5** (2), 79–141.

Gurhan-Canli, Z. and Maheswaran, D. (2000) Cultural variations in country of origin effects. *Journal of Marketing Research*, **37**, 309–317.

Keaveney, S.M. (1995) Customer switching behavior in service industries: an exploratory study. *Journal of Marketing*, **59** (2), 71–83.

Peterson, R.A. and Jolibert, A.J.P. (1995) A meta-analysis of country-of-origin effects. *Journal of International Business Studies*, **26** (4), 883–900.

Roggeveen, A.L., Bharadwaj, N., and Hoyer, W.D. (2007) How call center location impacts expectations of service from reputable versus lesser known firms. *Journal of Retailing*, **83** (4), 403–410.

Venkatraman, N.V. (2004) Offshoring without guilt. *Sloan Management Review*, **45** (3), 14–16.

Whoriskey, P. (2008) The Bangalore backlash: call centers return to U.S. *The Washington Post*, December 11, p. D01.

Zeithaml, V.A., Berry, L.L., and Parasuraman, A. (1996) The behavioral consequences of service quality. *Journal of Marketing*, **60** (2), 31–46.

marketing strategy implementation

Robert E. Morgan

Introduction

A critical indictment of contemporary marketing practice is that "misguided marketing strategies have destroyed more shareholder value, and probably more careers, than shoddy accounting or shady fiscal practices have" (McGovern *et al.*, 2004, p. 70). The object of such derision would appear to be the MARKETING STRATEGY per se but, all too frequently, this conveys the symptom rather than the cause of poor marketing practice. As White, Conant and Echambadi (2003) observed, marketing strategy implementation (MSI) capability perfectly mediates the relationship between marketing strategy development style and firm performance. Consequently, MSI has become an important theoretical focus for researchers, a key performance indicator for marketing personnel, and a vital precursor to organizational effectiveness.

Definition

Strategy implementation is traditionally considered as the "doing" stage in the MARKETING PLANNING process and follows the "dreaming" stage, commonly referred to as strategy formulation (Olson, Slater and Hult, 2005). Considered synonymous with marketing strategy execution, delivery, and transition, MSI can be defined as the *communication, interpretation, adoption, and enactment of a marketing strategy or strategic marketing initiative* (Noble and Mokwa, 1999, p. 57). Despite the fact that the normative literature has recently declared strategy implementation as the *sine qua non* of firm performance gains (Bower and Gilbert, 2007), the literature that underpins our academic knowledge of this concept reflects the embryonic stage of its theoretical development.

Scope of Thinking

Figure1 outlines the coverage of the strategy implementation literature from specifically the marketing domain.[1] This figure indicates the range of perspectives, influences, and foci that have been employed in studying the multifaceted and complex organizational process of MSI.

Historically, studies initially focused on structural phenomena in seeking to explain successful MSI; these studies commonly adopted either the structure-conduct-performance framework or institutional theory (Walker and Ruekert, 1987). This was followed by a wave of studies that examined the interpersonal processes that underlie the implementation effort (Noble, 1999). Such investigations focused primarily on the problems and challenges encountered in MSI with the conclusions being that poor marketing performance could be attributable to the disconnect between the formulation of marketing strategy and its implementation (Piercy, 1998). This dichotomy leads to a fault line in the process of strategizing whereby the harsh realities of implementation are misunderstood by the idealizing that can be characteristic of formulation stage of the planning process. At this juncture in the development of MSI knowledge, a variety of lenses were employed which ranged from contingency explanations through to psycho-social theories (Chimhanzi and Morgan, 2005) and more mainstream resource-based approaches (Hughes and Morgan, 2008). Consequently, our developed in different ways but much of the research conducted examinined the role of the individual/group or the behavioral and cultural processes that underpin MSI (Noble, 1999).

In a similar way, the level of analysis has changed over time. Initially focused on the top management team and chief marketing officer (Homburg, Workman and Krohmer, 1999), the role of the mid-level (marketing) manager became apparent as a key linchpin in delivering the implementation actions required of the marketing strategy (Thorpe and Morgan, 2007). More recently, the level of analysis has also included frontline employees and the role of the customer in delivering effective MSI for firms (Homburg, Wieseke and Bornemann, 2009).

The cascade of supporting actions and decisions required to realize marketing strategy means that there are inevitable implications at the tactical and operational level of the firm (Porter and Harper, 2003). Often overlooked, marketing actions at this level can have a seemingly disproportionate consequence on MSI.

Wiley International Encyclopedia of Marketing, edited by Jagdish N. Sheth and Naresh K. Malhotra.

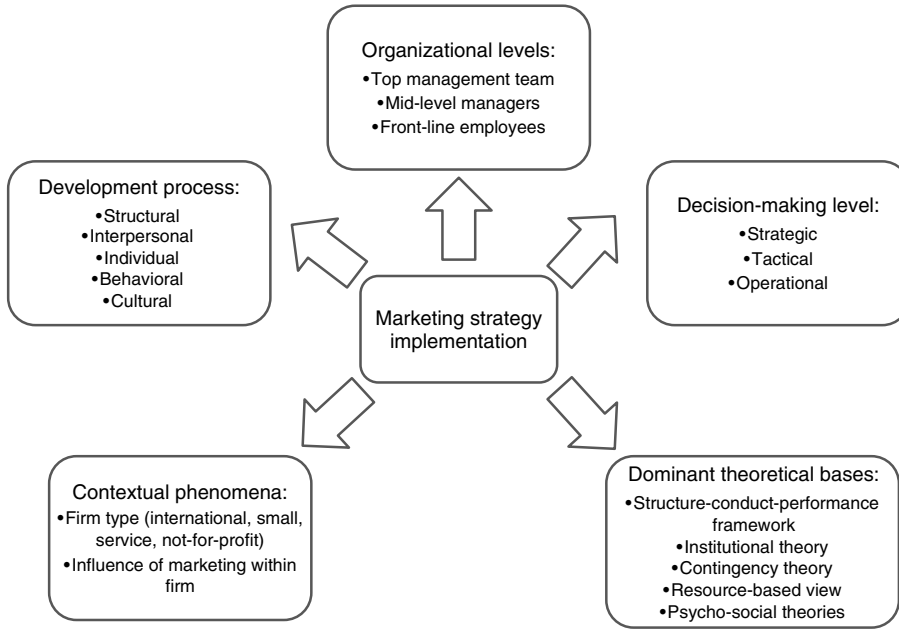


Figure 1 The main theoretical explanada underlying marketing strategy implementation.

Finally, MSI considerations remain as important for the multinational corporation as they do for other types of organizations from the small firm to the service-based enterprise and not-for-profit sector (Sashittal and Jassawalla, 2001).

Observations

There is overwhelming evidence to indicate that, provided marketing strategy is cogent, effective MSI can deliver performance gains for firms. Effective MSI derives from, among other factors, the appropriate organizational structures and administrative systems; control mechanisms; leadership style; employee incentives; communication efforts; integration and coordination arrangements; and, levels of strategy commitment and consensus. However, despite implementing a marketing strategy at one point in time, awareness needs to be given to the strategy dynamics at work which can significantly affect how optimum the strategy remains over time. A concept that is currently receiving considerably more attention by managers and researchers alike is strategic improvisation, which describes the simultaneous formulation and implementation of a strategy implying a responsiveness and flexibility in planning which is unparalleled.

Endnotes

[1] Much of the literature associated with the concept of strategy implementation derives from the organizational behavior, administrative science, public policy, technology, strategy and marketing disciplines. There is significant cross fertilization of theoretical ideas and applications among the literatures pertaining to implementation issues, processes, and contexts. This discussion is devoted specifically to sources within the marketing domain and observations that derive from empirical studies of marketing strategy.

Bibliography

Bower, J.L. and Gilbert, C.G. (2007) How managers' everyday decisions create or destroy your company's strategy. *Harvard Business Review*, **85**, 72–79.

Chimhanzi, J. and Morgan, R.E. (2005) Explanations from the marketing/human resources dyad for

marketing strategy implementation effectiveness in service firms. *Journal of Business Research*, **58** (6), 787–796.

Homburg, C., Wieseke, J. and Bornemann, T. (2009) Implementing the marketing concept at the employee-customer interface: the role of customer need knowledge. *Journal of Marketing*, **73**, 64–81.

Homburg, C., Workman, J.P. and Krohmer, H. (1999) Marketing's influence within the firm. *Journal of Marketing*, **63**, 1–17.

Hughes, P. and Morgan, R.E. (2008) Fitting strategic resources with product-market strategy: performance implications. *Journal of Business Research*, **61** (4), 323–331.

McGovern, G.J., Court, D., Quelch, J.A. and Crawford, B. (2004) Bringing customers back into the boardroom. *Harvard Business Review*, **82** (11), 70–80.

Noble, C.H. (1999) The eclectic roots of strategy implementation. *Journal of Business Research*, **45** (2), 119–135.

Noble, C.H. and Mokwa, M.P. (1999) Implementing marketing strategies: developing and testing a managerial theory. *Journal of Marketing*, **63**, 57–73.

Olson, E.M., Slater, S.F. and Hult, G.T.M. (2005) The importance of structure and process to strategy implementation. *Business Horizons*, **48**, 47–54.

Piercy, N.F. (1998) Marketing implementation: the implications of marketing paradigm weakness for the strategy execution process. *Journal of the Academy of Marketing Science*, **26** (3), 222–236.

Porter, T.W. and Harper, S.C. (2003) Tactical implementation: the devil is in the details. *Business Horizons*, **46**, 53–60.

Sashittal, H.C. and Jassawalla, A. (2001) Marketing implementation in smaller organizations: definition, framework, and propositional inventory. *Journal of the Academy of Marketing Science*, **29** (1), 50–69.

Thorpe, E.R. and Morgan, R.E. (2007) In pursuit of the 'ideal approach' to successful marketing strategy implementation. *European Journal of Marketing*, **41** (5/6), 659–677.

Walker, O.C. and Ruekert, R.W. (1987) Marketing's role in the implementation of business strategies: a critical review and conceptual framework. *Journal of Marketing*, **51**, 15–33.

White, J.C., Conant, J.S. and Echambadi, R. (2003) Marketing strategy development styles, implementation capability, and firm performance: investigating the curvilinear impact of multiple strategy-making styles. *Marketing Letters*, **14** (2), 111–124.