

Chapter 2

Explanatory and Predictive Models of Consumer Behavior*

John H. Roberts

*Australian Graduate School of Management, University of New South Wales, Kensington,
NSW, Australia 2033*

Gary L. Lilien

Pennsylvania State University, 113 BAB II, University Park, PA 16802, USA

The engineer builds on well-established laws of physics that have been derived by theoretical analysis and tested by empirical investigation. Similarly, marketing managers rely on consumer behavior models – models of how individual purchasing agents act in the marketplace – which are based on consumer behavior theory [e.g. Zaltman and Wallendorf, 1979] and then tested in the marketplace. Our objectives for this chapter are (a) to outline the elements of some basic consumer behavior models; (b) to indicate some of the ways those models have been extended; and (c) to illustrate how marketing models have been and can be applied to solve real management problems. We proceed within a framework that decomposes and categorizes consumer decision processes into a number of stages.

The chapter is designed to give quantitative modelers and management scientists unfamiliar with marketing an appreciation of the way in which models of consumer behavior are developed and used. It is also designed to provide a reference and teaching resource for marketing specialists. We assume the reader has a reasonable knowledge of and interest in the mathematics of the models, but has only modest understanding of the marketing models field.

1. An organizing framework for consumer behavior models

The consumer behavior model that we use in a given situation and the consumer behavior theory on which it is based depend on the objectives of the model-builder,

*Some of the material in this paper, prepared by the same authors, also appears in Chapter 2 of Lilien, Kotler & Moorthy [1992], *Marketing Models* (Prentice-Hall, Englewood Cliffs).

the important market phenomena, and the availability of relevant theories and data to support the analysis. The range of consumer behavior models is very broad for several reasons:

Consumers are different. Consumers vary according to their personalities, values, preferences, and a range of other characteristics. These differences mean that a model that is appropriate for describing the behavior of one particular consumer may be inadequate in explaining the behavior of another, even in similar purchase situations.

Choice decisions differ. Not only do consumers differ from one another, but even for one specific consumer, a model that might describe that consumer's behavior for one specific purchase decision may not work for another product, for several reasons. For example, the level of involvement that the consumer has in the decision will determine the amount of cognitive effort and search that he or she is prepared to invest in it. Phenomena associated with low involvement decisions (often called routinized response behavior) include habit, indifference to risk, and lack of search. As the level of involvement grows through limited problem solving to extended problem solving, other phenomena such as brand perceptions and evaluation rules become important.

The context of purchases differs. Consumers vary in their decision-making rules because of the usage situation, the user of the good or service (for family, for gift, for self), and purchase situation (catalogue sale, in-store shelf selection, salesperson-aided purchase). Each such context may invoke a different decision-making strategy. For example, a consumer buying a watch for herself may value consistency with her other jewelry, image, and aesthetics, but when buying one as a gift for a friend, may treat price and manufacturer reputation as the most important attributes.

Managerial needs differ. Models are also important managerial tools and differences in managerial problems may also make some models more appropriate than others. For example, if management is interested in a pricing strategy, then a model that emphasizes the consumer's evaluation process and the role of price in that process may be most appropriate; if management is interested in making consumers aware of a new product launch, then a model that focuses on information search and perception formation may be best. A model that deals with all aspects of consumer behavior in complete detail may be theoretically sound but hopelessly complex in terms of its data requirements and potential for calibration.

Before proceeding to our review, we would like to indicate several categories of consumer behavior models that we explicitly exclude.

First, note that the models that we review generally assume that a single individual is acting as his or her own agent or, at most, as an agent for a household. Thus we exclude *organizational buying* models from consideration. Organizational buying is characterized by derived demand (products purchased by organizations as inputs to satisfy the needs of those organizations' customers), multiple individuals involved in the purchase process (the *buying center*), and extensive use of bargaining, negotiations and long-term relationships to effect an exchange [Lilien, Kotler & Moorthy, 1992, Chapter 3].

Second, we exclude so-called *stochastic models* from consideration. Traditionally, stochastic models have been suggested as modeling frameworks primarily for low-involvement products where little conscious decision-making was assumed to take place. Those models have generally focused most of their attention on the mechanics of an underlying probabilistic process [Massy, Montgomery & Morrison, 1970; Lilien, Kotler & Moorthy, 1992, Chapter 2]. In contrast, the models that we focus on here devote most of their attention to the determinants of choice, with uncertainty arising because of missing variables, simplified specification, incompleteness of consumer information, measurement error and the like. These latter models are generally associated with high-involvement purchase situations, i.e. those that consumers find personally significant. We recognize that there is an imperfect correspondence between stochastic models and low-involvement choice situations, as well as between explanatory/predictive models and high-involvement buying.

Finally, consumer behavior models can either deal with consumers as individuals or as an aggregate group. We deal here with models of the individual, although often with a view toward aggregating those models to make statements about the behavior of the total market. Other chapters in this volume deal with market response as a whole, looking at the relationship between market share and sales to marketing activities (promotion and price, for example) relying solely on aggregate, market-level data. An area of continuing research is the relationship between individual response models and aggregate market models [see, for example, Ehrenberg, Goodhardt & Barwise, 1990, and Chatterjee & Eliashberg, 1990].

2. A taxonomy of consumer behavior models

To achieve parsimony and to help structure our review, we classify consumer behavior models in the framework of a staged or phased process. That is, we visualize consumers going through a number of steps from the time that they recognize a need that they would like to satisfy, through the choice of a product to satisfy that need, through the actual purchase and consumption experience and, finally, through the updating of preferences and perceptions that follow consumption and guide future purchase behavior. Table 2.1 outlines the five organizing phases we use here.

We also use Table 2.1 to demonstrate the applications different marketing models have seen. The applications of these models will often cross over the boundaries of this framework. Although we realize that this classification is not definitive, it has helped us organize the literature and we hope it will help the reader understand the literature better.

The next five sections correspond to the basic elements in Table 2.1. In each of those sections we first review some basic concepts and models of that consumer behavior stage. Then we outline some important extensions and applications. These applications illustrate the types of extensions researchers have undertaken; they are not meant to be comprehensive. Where possible, we cite more complete reviews.

Table 2.1
A framework for classifying consumer behavior models

Stage	Dependent variables of interest	Typical models used for this stage
Need arousal	Purchase (category choice) Purchase timing (Table 2.2)	Binary choice models
Information search	Awareness (aided/unaided) Consideration/evoked set Choice set Belief dynamics (Table 2.3)	Individual awareness models Consideration models Information integration models
Evaluation	Product perceptions (Table 2.4) Product preferences (Table 2.5)	Perceptual mapping/ multidimensional scaling models Attitude models: Compensatory Non-compensatory
Purchase	Brand choice Store choice Quantity choice (Table 2.6)	Discrete choice models Hierarchical models
Post-purchase	Brand satisfaction/satiation Word-of-mouth (Table 2.7)	Satisfaction models Variety-seeking models Communications/network models

Some models span several stages of the taxonomy; we deal with a sample of such models in Section 8, illustrating how models of several of these stages can be combined to address more complex managerial problems. We conclude with a discussion of what we see as key future directions for consumer behavior models.

The elements of Table 2.1 correspond to the following basic consumer processes:

(1) *Need arousal*. A need can be activated (or aroused) through internal or external stimuli. In the first case, one of the person's normal drives – hunger, thirst, sex – rises to a threshold level and becomes a drive. In the second case, need is caused by an external stimulus or triggering cue (an advertisement, sight of an acquaintance's product, etc.). Consumers then determine the type of product that could possibly satisfy the need.

(2) *Information search*. Consumers often do not satisfy an aroused need immediately. Depending on the intensity of the stored need, consumers either undertake *active information search*, or enter a state of *heightened attention* in which they are open to the reception of information passively.

Following the search process, consumers are aware of a group of products or brands that they see as being possibly suitable to satisfy the identified need. This

group of products is called the *evoked set*, or *consideration set*. The consideration set comprises the alternatives that enter the next stage, the evaluation phase. As evaluation progresses, further brands may be eliminated: the consideration set is dynamic. When consumers are about to make a purchase, the remaining set of brands is the *choice set*.

(3) *Evaluation*. Evaluation has two components. Consumers establish their beliefs about the features of the alternative products that they consider (*perceptions*) and they determine, based on those perceptions, their attitudes towards the products (*preferences*). In practice, consumers update perceptions continually during the search process.

Perception formation. The fields of psychology [Fishbein, 1967] and economics [Lancaster, 1966] both suggest that consumers see a product as having several attributes. Consumers perceive a particular product in terms of where it lies in the space spanned by the set of attributes relevant to its product class. For example, in the aspirin category, important attributes might be speed of relief, reliability, side effects and price. Individual consumers vary as to which attributes they consider most relevant.

Consumers are likely to develop opinions about where different brands stand on each attribute. The set of beliefs that consumers hold about a particular brand is known as its *brand image*. Consumers' beliefs or perceptions may differ from the 'true' attributes because of consumers' particular experiences and the way consumers gather and process information. It is consumers' perceptions of a product's characteristics that influence their behavior, not the 'true' characteristics.

Preference formation. Consumers use their perceptions in forming brand preferences. Most models assume consumers have a utility function for attributes that describes how the consumer's valuation of the product varies with alternative levels and combinations of attributes. Consumers arrive at an attitude (judgment or preference) toward the brand through some evaluation procedure. Starting with their consideration set, consumers compare products and end up with an ordering of preferences (although not all evaluations follow this process).

(4) *Purchase decision*. When evaluating products, consumers may form a ranked set of preferences for the alternative products in their consideration sets and develop an intention to purchase the product they like best. But a number of additional factors often intervene before a purchase can be made [Sheth, 1974]. One factor is the attitude of others, including the intensity of others' attitudes, and the consumers' motivation to comply with others' wishes [Fishbein, 1967]. Consumers' purchase intentions are also influenced by changes in anticipated situational factors, such as expected family income, the expected total cost of the product, and the expected benefits of the product. Furthermore, when consumers are about to act, unanticipated situational factors may intervene to prevent them from doing so (such as the lack of availability of a preferred product). Finally, measurement error may arise when we try to estimate preferences. Thus, estimated preferences and purchase intentions are not completely reliable predictors of actual buying behavior: while they indicate likely purchase behavior, they fail to include a number of additional factors that may intervene.

(5) *Post-purchase feelings.* After buying and trying the product, consumers will experience some level of satisfaction or dissatisfaction. Swan & Combs [1976] posit that consumers' satisfaction is a function both of expectations and the product's perceived performance. If the seller makes exaggerated claims for the product, consumers experience disconfirmed expectations, which lead to dissatisfaction. The level of satisfaction or dissatisfaction depends on the size of the difference between expectations and performance. Satisfaction with a product will influence consumer choice on subsequent purchase occasions. In addition, consumers are likely to communicate their feelings about the product to other potential consumers who are seeking information. Satisfaction is a more powerful influence in frequently purchased goods where the purchaser's own experience is critical for repurchase and repurchase rates are high, while the opinion of others is a more important consideration for durable products.

We now deal with these elements in our taxonomy in more detail.

3. Need arousal

Need arousal is the trigger that starts the consumer decision process. The modeling of whether and when that need will be satisfied, in our view, corresponds to a category purchase decision. Later we will review *which* specific product or brand the consumer chooses. The models at both stages are similar: the choice of whether to buy or not and the choice of what to buy. Models of these processes draw on discrete choice theory.

When there are exactly two choices (buy in category/don't buy in category, as here) discrete choice models are called *binary choice models*, and have been applied to a wide variety of classification problems within marketing and in other areas [Ben-Akiva & Lerman, 1985].

3.1. Need arousal basics

Assume that the utility that consumer i expects to get from the category at the time of the purchase decision is U_{Bi} , while the utility of not buying within the category is U_{Ni} . Furthermore, assume that we can divide these utilities into two components; a systematic part, V_i and a random component ε_i .

Thus,

$$U_{Bi} = V_{Bi} + \varepsilon_{Bi}, \quad \text{and} \quad U_{Ni} = V_{Ni} + \varepsilon_{Ni}, \quad (1a, b)$$

or, i.e.,

$$\text{Buy/Not Buy Utility} = \text{True Value} + \text{Assessment Error.}$$

(In what follows, we drop the subscript i for simplicity.)

There are two components in Equation (1): the true utility value (V) and the assessment error component (ε). One way of structuring V is to compare the utility

(net of price) that the category gives to the utility of other alternative uses of that amount of money (the budget constraint). Hauser & Urban [1986] derive and test such a rule, described below. Thus, binary choice models vary in the benchmark used for not buying in the category (V_N). They also vary in their assumptions about the distribution of the disturbance term, ε .

Assuming the following:

- (1) $\{\varepsilon_B\}$ and $\{\varepsilon_N\}$ are independent and identically distributed; and
- (2) the distribution of $\{\varepsilon_B, \varepsilon_N\}$ is double-exponential (extreme-value):

$$\Pr(\varepsilon \leq x) = \exp\{-e^{-bx}\}.$$

Then we get:

$$\Pr(\varepsilon_N < X_N) = \exp\{-e^{-bX_N}\}. \quad (2)$$

To buy in the category requires that

$$U_B > U_N$$

which occurs with probability:

$$\Pr(\varepsilon_N < V_B + \varepsilon_B - V_N). \quad (3)$$

To evaluate (3), then, we substitute $V_B + \varepsilon_B - V_N$ for X_N in (2):

$$\Pr(\varepsilon_N < V_B + \varepsilon_B - V_N) = \exp\{-e^{-b(V_B + \varepsilon_B - V_N)}\} \quad (4)$$

We must now integrate ε_B (the remaining random variable) out of (4):

$$\Pr(\text{Buy}) = \int_{-\infty}^{\infty} \exp\{-e^{-b(V_B + \varepsilon_B - V_N)}\} f_{\varepsilon_B}(\varepsilon_B) d\varepsilon_B \quad (5)$$

where

$$f_{\varepsilon_B}(\varepsilon_B) = b \exp\{-b\varepsilon_B\} \exp\{-e^{-b\varepsilon_B}\}. \quad (6)$$

After some algebra, the logit model evolves:

$$\Pr(U_B > U_N) = \frac{1}{1 + e^{-b(V_B - V_N)}} \quad (7a)$$

or, if the benchmark for not buying is $V_N = 0$, then

$$\Pr(U_B > U_N) = \frac{1}{1 + e^{-bV_B}}. \quad (7b)$$

3.2. Need arousal extensions and applications

Two issues that distinguish need arousal models are their specifications of the systematic component and the specifications of the error term (Table 2.2).

3.2.1. Specification of the systematic components: V_B and V_N

A key question that arises in specifying the systematic component in Equation (1b) is how to determine the utility of not buying within the category V_{Ni} . Various authors have suggested different benchmarks against which the utility of buying within the category can be compared. For example, budget constraints compare the utility (net of price) that the category in question gives to that of other durables and a composite good of non-durables. This will result in an ordering that consumers can use to purchase durable products.

Example. Hauser & Urban [1986] derive and test a decision rule where they posit that consumers undertake this budgeting process, called the value priority algorithm. Under the value priority algorithm, consumers select the durables of highest utility per dollar, or highest utility net of price, first and proceed down the list ordered on that basis until they meet their budget constraint. That problem may be expressed as the following linear program (assuming linear, additive utilities of durables):

$$\begin{aligned} &\text{maximize} && u_1g_1 + u_2g_2 + \dots + u_n g_n + u_y(y) \\ &\text{subject to} && p_1g_1 + p_2g_2 + \dots + p_n g_n + y \leq B, \\ &&& g_j \geq 0 \quad \text{for all durables } j \end{aligned}$$

where

$$\begin{aligned} u_j &= \text{expected utility of durable } j, \\ p_j &= \text{price of durable } j, \\ g_j &= \begin{cases} 1 & \text{if durable } j \text{ is purchased and} \\ 0 & \text{otherwise,} \end{cases} \\ B &= \text{budget that the consumer has to spend, of which...} \\ y &= \text{money spent on products other than durables, giving a utility of } u_y(y). \end{aligned}$$

An equivalent problem for the consumer is to minimize the dual linear program, that is to solve the following problem:

$$\begin{aligned} &\text{minimize} && B\lambda + \gamma_1 + \gamma_2 + \dots + \gamma_n \\ &\text{subject to} && \gamma_j \geq u_j - \lambda p_j \quad \text{for all } j, \\ &&& \lambda = \partial u_y(y) / \partial y. \end{aligned}$$

The behavioral interpretation of this problem is that γ_j is the shadow price of the constraint $g_j \leq 1$. That is, γ_j is the forgone value of not having durable j or the value of relaxing the constraint that durables are discrete. By the rule of

Table 2.2
Need arousal models: extensions and applications

Issues addressed	Model	Data	Illustrative literature	Comments/application
Alternative benchmarks of deterministic term:				
Other durables	Value priority (math. programming)	Category purchase intent plus other durables Utility and prices of category and other durables	Hauser & Urban [1986]	Durable good benchmark
Current holdings	Existing stock model	Category purchase intent Utility of new purchase plus utility of existing stock	Hauser, Roberts & Urban [1983]	Currently held durable benchmark
Time-varying utility	Hazard rate models	Purchase intent or purchase level, dynamics and determinants of hazard	Jain & Vilcassim [1991]	Emphasis on dynamics of probability
Alternative specification of error term				
	Binary logit	Category purchase or intent Utility of category Utility of alternative	Domencich & McFadden [1975] Bodnar, Dilworth & Iacono [1988] Robinson [1986] Nooteboom [1989] Lisco [1967]	Urban travel
	Binary probit	As for binary logit	Amemiya [1981]	Transport choice
Overview articles				
	Review of discrete choice models Diagnostics and comparison		Gessner, Kamakura, Malhotra & Zmijewski [1988]	

complementary slackness, net utility or consumer surplus, $u_j - \lambda p_j$, is greater than zero if and only if durable j is purchased.

Hauser & Urban fit the model by combining four measures of purchase intent: the reservation price of the durable, the stated probability of purchase, and two different orders of preference for the durable in a lottery. Both value per dollar and net value (assuming the consumer maximizes net utility) predict individual budget plans adequately for the majority of consumers: 60% have a correlation between predicted and actual plans of greater than 0.5 with the former and 84% for the latter.

In addition to other durable goods, benchmarks against which the new purchase might be compared include the utility of existing stock in consumer marketing [Hauser, Roberts & Urban, 1983] and the utility that purchase of the product will provide in business to business marketing (the present value of future income streams [Mansfield, 1961]). Alternatively, the utility of the category after adjusting for price may be compared to the utility of not having purchased, i.e. zero. If the net category utility exceeds zero, the product will be purchased and is said to offer a consumer surplus to the purchaser equal to its net utility. Other, related recent work has been emerging using hazard functions to model the purchase timing decision [Jain & Vilcassim, 1991].

Having specified the expected utility of not buying within the category, V_N , let us consider the determinants of the expected utility of buying within the category, V_B . This can be done in two ways. First, buying within the category can be characterized by the utility that a *typical* product would offer in terms of its anticipated attribute levels. Alternatively, consumers might estimate their expected utilities of all the brands that they would consider. From the utilities of the individual brands, they can gain an estimate of the utility of the category as a whole. For example, if a logit model were used to describe brand choice, then the expected utility of the category would be

$$V_B = E(\max(U_j)) = \ln \left(\sum_{j \in C} \exp V_j \right) \quad (8)$$

where V_j is their expected utility for the j th brand considered (in set C) [Ben-Akiva & Lerman, 1985, p. 105].

3.2.2. Specification of the error component

We considered an extreme value distribution for the error earlier. The most common alternative is to treat the error term as the sum of a large number of unobserved influences. If we can assume independence of those influences then the sum of all of these factors will tend to be normal by the central limit theorem. If we assume that $(\varepsilon_N - \varepsilon_B)$ is distributed as $N(0, \sigma^2)$ (with the corresponding assumed normality but not necessarily independence of ε_N and ε_B), then we can derive the probability of consumer i buying within the category, P_B as a function of the

expected utility components, V_B and V_N :

$$P_B = \Pr((\varepsilon_N - \varepsilon_B) > V_B - V_N) \quad (9a)$$

$$= \int_{-\infty}^{V_B - V_N} 1/(\sigma\sqrt{2\pi}) \exp(-(x/\sigma)^2/2) dx. \quad (9b)$$

Equation (9b) is called the binary probit model.

3.2.3. Other issues in binary choice

There are a number of other binary choice models available, including the linear probability model (assuming a uniform distribution of ε) and arctan (assuming an arctan distribution) [see Ben-Akiva & Lerman, 1985]. Gessner, Kamakura, Malhotra & Zmijewski [1988] reviewed the relative performance of these and other binary choice models and suggest that in the absence of major violations to their assumptions, all of the choice models they examined (probit, logit, linear probability and two types of discriminant function) fit and predict reasonably well, giving qualitatively similar results. However, this is not the case if there are major violations to the assumptions. Gessner et al. conclude that the choice of model should be data-dependent, that there is little difference between the binary logit and probit models, and that users of these models should test for data inadequacies and transform their data where necessary to avoid them.

4. Information search

Once a consumer recognizes a need, he enters a state of heightened awareness in which he seeks more information about brands or products that could satisfy that need. Evaluation and brand choice take place based on the information resulting from this search.

Information search has two parts. The first is understanding the status of each brand during the search: Will the consumer be aware of it? Will he consider it worth searching further? The second part is how information discovered during search is incorporated into the consumer's belief structure – that is, the dynamics of evaluation beliefs as a result of new information. In practice, the consumer is likely to have prior beliefs about the alternatives that could satisfy his need. The structure of those beliefs and how they are combined to form preferences are discussed in the next section.

Thus, the basic tools for understanding information search are:

- Will awareness be obtained?
- Will aware brands be considered?
- How does the information gained during search affect beliefs?
- When does the consumer stop searching?

4.1. Information search basics

4.1.1. Models of awareness

Rossiter & Percy [1987] define awareness as 'the buyer's ability to identify (recognize or recall) the brand within the category in sufficient detail to make a purchase'. Most of the related literature links aggregate levels of awareness to levels of advertising and most models have no individual-level interpretation [see Mahajan, Muller & Sharma, 1984]. An exception is work by Blattberg & Jeuland [1981].

There are two elements of individual-level awareness models: advertising exposure and forgetting. Blattberg & Jeuland [1981] use a Bernoulli advertising exposure process and an exponential forgetting process to model awareness. The Bernoulli assumption implies that if there are n advertisements during a period, then the probability that a consumer will be exposed x times is

$$\Pr(x \text{ exposures}) = [n!/(x!(n-x)!)]q^x(1-q)^{n-x} \quad (10)$$

where q is a parameter. Exponential forgetting suggests that if the last advertisement was seen by the consumer at time t_1 , the probability of him remembering it (still being aware) at time t , p_t , is given by:

$$p_t = \exp\{-\alpha(t-t_1)\} \quad (11)$$

where α is a parameter (the forgetting rate).

The probability of a consumer being aware at time t , $f(t)$, may be calculated in terms of the probability of the consumer having seen the most recent advertisement (at time t_1) times the probability of not having forgotten it, plus the probability of having seen the previous advertisement (at time t_2) and not forgetting that (given that he did not see the most recent advertisement), etc. Mathematically we may write:

$$\begin{aligned} f(t) &= q \exp\{-\alpha(t-t_1)\} + q(1-q) \exp\{-\alpha(t-t_2)\} \\ &\quad + q(1-q)^2 \exp\{-\alpha(t-t_3)\} + \dots \\ &= \sum_n q(1-q)^{n-1} \exp\{-\alpha(t-t_n)\}. \end{aligned} \quad (12)$$

At the aggregate level, the interpretation of $f(t)$ is the expected proportion of the target population that is aware.

4.1.2. Models of consideration-set formation

Many empirical studies show that consumers do not search and evaluate (consider) all the brands of which they are aware. A key question that arises in modeling consideration is whether the process should be compensatory (in which shortcomings in one attribute may be traded off against benefits on another) or non-compensatory (in which certain thresholds exist for different attributes and

the brand must meet some combination of those thresholds, irrespective of its levels of other attributes). While evidence exists for both depending on the situation, in this section we outline a basic compensatory model. Models of non-compensatory processes are described in Section 5.

If we assume that the consumer will choose from the consideration set according to the logit choice model at the purchase stage, as described in Section 6, then (using Equation (8)), we can estimate the expected utility that he or she will derive from buying within the category, given a consideration set of C , $E_{B|C}$ [see, for example, Roberts & Lattin, 1991]:

$$E_{B|C} = \ln\left(\sum_{j \in C} \exp(U_j)\right). \quad (13)$$

If the consumer now becomes aware of a new brand, N , with search costs c_N and utility U_N we can use Equation (13) to estimate whether or not it will be considered. It should be considered if the incremental expected benefit from the new consideration set $E_{B|C \cup N}$ more than offsets the cost of search, c_N ; that is, if

$$E_{B|C \cup N} - E_{B|C} > c_N. \quad (14)$$

Substituting the expression for expected category utility from Equation (13), and rearranging terms, we can derive the minimum utility that the brand needs to justify entry into the set, or alternatively, the maximum search costs that it can sustain to be included:

$$E_{B|C \cup N} - E_{B|C} > c_N \quad \text{if} \quad \ln\left(\sum_{j \in C \cup N} \exp(U_j)\right) - \ln\left(\sum_{j \in C} \exp(U_j)\right) > c_N,$$

i.e.

$$U_N > \ln\left\{\left[\sum_{k \in C} \exp(U_k)\right][\exp(c_N) - 1]\right\} \quad (15a)$$

or

$$c_N < \ln\left\{1 + \exp(U_N) / \left[\sum_{k \in C} \exp(U_k)\right]\right\}. \quad (15b)$$

Equation (15b) shows that even if all brands are of equal utility ($U_k = U$), as the number of brands already considered increases, the maximum search costs that an additional brand can justify decreases.

4.1.3. Information integration

The consumer searches those brands he is aware of and which merit consideration. Basic models address how information discovered during search should be

integrated into consumers' perceptions and whether more search should be conducted. The most commonly used updating procedure follows Bayes's rule. If prior beliefs about attribute k are normally distributed with mean y_k and variance σ_k^2 and new information that is received is also distributed normally (mean \bar{y}_k and variance σ_y^2) then beliefs after updating will be normally distributed with mean $y_{k'}$ and variance $\sigma_{k'}^2$ [DeGroot, 1970] where

$$y_{k'} = (\sigma_y^2 y_k + \sigma_k^2 \bar{y}_k) / (\sigma_y^2 + \sigma_k^2) \quad \text{and} \quad (16a)$$

$$\begin{aligned} \sigma_{k'}^2 &= \{\sigma_y^2 / (\sigma_y^2 + \sigma_k^2)\}^2 \sigma_k^2 + \{\sigma_k^2 / (\sigma_y^2 + \sigma_k^2)\}^2 \sigma_y^2 \\ &= \sigma_k^2 \sigma_y^2 / \{\sigma_y^2 + \sigma_k^2\}. \end{aligned} \quad (16b)$$

The prior and sampling distributions are called a normal-normal conjugate pair [e.g. see Roberts & Urban, 1988].

4.1.4. Optimal stopping rules

Hagerty & Aaker [1984] have developed a model for information search strategies based on the sequential sampling literature that looks at the expected value of sample information (EVSI). They assume a utility-maximizing consumer who will choose the piece of information with the greatest difference between the expected value of the information to search next and the information processing cost. This approach builds on the economics of information literature developed by Stigler [1961] and applied in marketing by Shugan [1980].

To illustrate how the expected value of the sample information (EVSI) is calculated, assume that the consumer is currently considering three brands, 1, 2 and 3, ordered in terms of decreasing expected utility, $E(U_j)$. There is some uncertainty associated with each brand j , σ_j^2 . If the distribution of beliefs about the utility of each brand is distributed normally, and if the consumer does not gather any more information, he will choose brand 1, since it has the highest expected utility. To estimate the value of additional search of brand 1 consider the expected utility of brand 1, after the new information has been gathered.

Search will only change the consumer's choice and thus alter the utility that he realizes if, based on the new information, brand 1 has an updated expected utility, m , of less than $E(U_2)$. In that case the expected utility that would be forgone by not undertaking the search is $E(U_2) - m$. What search enables the consumer to do is to reduce the chances of incurring that loss by giving him a better fix on the true value of $E(U_1)$. The expected value of search (EVSI) is $E(U_2) - m$ integrated over the probability distribution of different values of m , $p(m)$, from $-\infty$ to $E(U_2)$:

$$\text{EVSI} = \int_{-\infty}^{E(U_2)} (E(U_2) - m)p(m)dm. \quad (17)$$

Since prior beliefs are normally distributed, if we assume that new information is normally distributed, the posterior mean (the utility of brand 1 after updating,

m) will also be normally distributed, with variance σ_m^2 . From this, Hagerty & Aaker show that in the multibrand case the expected value of sample information from searching brand j may be rewritten:

$$\text{EVSI} = \sigma_m \Phi(\delta_y / \sigma_m) \quad (18)$$

where

$$\delta_y = \begin{cases} E(U_1) - E(U_j) & \text{for } j \neq 1, \\ E(U_1) - E(U_2) & \text{for } j = 1, \end{cases}$$

Φ = the integral of the standard normal distribution.

The extension of this formula to searching all brands is analogous to Equation (18). The problem the consumer faces in deciding which piece of information, y , to process at time t then becomes:

$$\begin{aligned} &\text{maximum} \quad \text{EVSI}_t^y - c^y \\ &\text{subject to} \quad \text{EVSI}_t^y - c^y > D \geq 0. \end{aligned} \quad (19)$$

That is, the consumer must maximize the value of the next piece of information minus processing costs as long as the search has a value of greater than D , a threshold below which it is not worth the effort of searching.

4.2. Information search models: extensions and applications

Table 2.3 outlines some key extensions and applications of information search models.

4.2.1. Awareness

Leckenby & Kishi [1984] derive the Dirichlet multinomial distribution (DMD) to model the proportion of the population that will be exposed to $1, 2, 3, \dots, mn$ of the n insertions in each of m media vehicles. Their model comes from the assumption that exposure to the mn advertisements for any individual consumer is multinomial, while individuals' parameters are distributed Dirichlet across the population. Leckenby & Kishi find superior performance of this model (estimated using a variety of algorithms) over other similar models.

4.2.2. Consideration-set formation

In an alternative approach to modeling consideration-set composition, Hauser & Wernerfelt [1990] assume a probit model at the choice stage rather than logit. They distinguish between the cost of search and the cost of evaluating and deciding between brands. They show that low-search-cost brands are more likely to be considered.

Large consideration sets for individual consumers are associated with lower search and evaluation costs and a higher variance of brands' utilities across

Table 2.3
Information search models: extensions and applications.

Issues addressed	Model	Data	Illustrative literature	Comments/application
Awareness creation: Dynamics of awareness	Bernoulli exposure and exponential forgetting	Advertising levels Awareness or sales at an individual or aggregate level	Blattberg & Jeuland [1981]	Fitted at aggregate level
Variety of advertising media	Dinichlet exposure	Advertising schedule	Leckenby & Kishi [1984]	Fitted at aggregate level
Consideration-set formation: Consideration-set membership	Cost-benefit inclusion: probit choice	Consideration, utility	Hauser & Wernerfelt [1990]	Aggregate testing
Consideration-set formation and size	Cost-benefit inclusion: logit choice	Consideration, utility, perceptions	Roberts & Lattin [1991]	Existing brands and new concepts
Fuzzy consideration sets	Probabilistic consideration	Probability of consideration, choice	Fotheringham [1988]	No empirical testing
Information integration: Information search	Hedonic pricing	Market attributes and prices	Ratchford [1982]	Benefits of search
Information integration	Expected value of sample information	Attribute importances and search behavior	Hagerty & Aaker [1984]	Optimal search patterns
	Linear updating	Prior beliefs and uncertainties	Meyer & Sathi [1985]	Linear updating
	Bayesian updating	Posterior beliefs and uncertainties	Roberts & Urban [1988] Oren & Schwartz [1988]	Normal conjugate pairs Beta binomial pairs
Overview	Review of awareness models	Strength and value of word-of-mouth	Mahajan, Muller & Sharma [1984]	Emphasis on aggregate models

consumption occasions. To establish the market-level implications of consideration, Hauser & Wernerfelt concentrate on testing their model at the aggregate level. They show that observed distributions of consideration-set sizes, order-of-entry effects, asymmetric advertising effects, and level of promotional activity in a number of packaged goods markets are consistent with their model.

In addition to industrial and consumer products, the concept of consideration has also been applied to the selection of retail outlets. Fotheringham [1988] suggests that consideration sets are not dichotomous, but rather are fuzzy sets (whether because consideration is not a discrete process to consumers or because as researchers we are not capable of measuring it). Thus he suggests that brands will be taken into the evaluation and choice phase only with some probability. For the problem of retail choice he makes that probability a function of the closeness of other stores in the awareness set. If closeness leads to a positive effect on the probability, he terms this an agglomeration effect. If it decreases the probability, he calls that a competitive effect.

Application of consideration models has increased in recent years, both for durables and packaged goods. It is important to note that the consideration set is dynamic and if consideration sets are calibrated well before the consumer enters the need-arousal stage, then the results may not be good indicators of later behavior [Day & Deutscher, 1982].

4.2.3. *Information search and integration*

Updated perceptions are generally a weighted average of consumers' beliefs and the sample information discovered [e.g. Meyer & Sathi, 1985]. Bayes rule, outlined in Equations (16a) and (16b) is the most common form of weighting. In addition to the normal-normal conjugate pairs illustrated there, other conjugate pairs of distributions include the beta-Bernoulli [e.g. see Oren & Schwartz, 1988], gamma-Poisson and exponential. Studies testing the predictive performance of Bayesian updating in consumer decision processes have had mixed findings [see Roberts & Urban, 1988, for a review]. However, in the absence of an obvious alternative contender, Bayes's rule has had considerable popularity in dynamic beliefs updating models.

5. Evaluation

Our taxonomy considers product evaluations as involving two processes: product perceptions and the relationship of those perceptions to preferences.

5.1. *Perceptual measurement basics*

Beliefs about products (perceptions) can be measured directly by asking consumers how much of a feature they perceive a certain product to contain, or they can be inferred, by asking consumers how similar certain products are and then inferring what discriminates between different products.

The two analytical approaches most frequently used to derive evaluation criteria and build perceptual maps are decompositional methods, based on multidimensional scaling (MDS), and compositional methods, based on factor analysis (FA). MDS procedures infer dimensions that discriminate between consumers' evaluations of different products based on brand interrelationships, while FA methods take explicit attribute data and distill them into underlying dimensions or factors.

5.1.1. Multidimensional scaling (MDS)

MDS is a set of procedures in which a reduced space depicting product alternatives reflects perceived similarities and dissimilarities between products by the inter-product distances. Different types of multidimensional scaling may be distinguished on the basis of the type of data input to the model, the number of dimensions on which the data are collected (modes), and the geometric model used to analyze the data (see DeSarbo, Manrai & Manrai (Chapter 5) and Lilien, Kotler & Moorthy [1992]). Nonmetric multidimensional scaling techniques are applied where similarities have ordinal scaling while metric methods are used with interval-scaled data.

The idea behind MDS is to create a map representing the product stimuli or consumer preferences. The objective is to have the interproduct distances in the map have the same rank order as the direct similarity judgments of products or preferences.

Let δ_{ij} denote the perceived dissimilarity between product alternatives i and j , which can either be obtained directly or be derived from distances using attribute rating scales. Then, with MDS, we find a configuration of points (the product alternatives) in a space of lowest dimensionality such that the ranking of inter-point distances d_{ij} is as close as possible to the ranking of the original dissimilarities δ_{ij} . This result is a monotonic relationship between the d_{ij} 's and the δ_{ij} 's. To reach this objective, MDS algorithms minimize a quantity called stress:

$$\text{Stress} = \left[\frac{\sum_{i < j} (\hat{d}_{ij} - d_{ij})^2}{\sum_{i < j} d_{ij}^2} \right]^{1/2} \quad (20)$$

where \hat{d}_{ij} is a distance as close as possible to the d_{ij} but is monotonic with the original dissimilarities δ_{ij} .

For a given dimensionality the configuration retained is the one that minimizes the stress function. The resultant map shows the relationship between the various products in the market. It may be arbitrarily reflected or rotated to aid interpretability.

This concept may be extended to incorporate both products and preferences on the same map, called a joint space. The process of deriving these two joint spaces is called *unfolding*.

There are also a number of other ways in which joint spaces can be presented. Brands can be represented as points, while respondents' preferences may be represented as vectors. This is called a *projection model* and is more appropriate when the dimensions are monotonically increasing in preference ('more is better' for each dimension). An individual's preference for brands can be obtained by pro-

jecting the brand onto the individual's ideal vector. (For a typical application see Moore & Winer [1987].)

Issues that must be considered with multidimensional scaling include the number of products needed, the determination of the dimensions, and the validity of the process. Green & Wind [1973] suggest that the number of dimensions should be less than one-third of the number of products. In practice, the consideration-set size provides an upper bound on the number of brands that can be evaluated. The stress measure can help determine the number of dimensions while the naming of the dimensions can be aided by examining their correlation with brand attribute ratings (if available). Tests of the reliability and validity of MDS have produced encouraging results and the methods are reasonably robust with respect to measurement error.

5.1.2. Factor analysis

Factor analysis was originally developed in connection with efforts to identify the major factors making up human intelligence. Since then, it has been applied to many other problems and is a frequently used technique in performing product-evaluation analyses in marketing.

The basic factor analysis model assumes that original perceptual ratings about a product are generated by a small number of latent variables, or factors, and that the variance observed in each original perceptual variable is accounted for partly by a set of common factors and partly by a factor specific to that variable. The common factors account for the correlations observed among the original variables. This model can be written as

$$x_{ijk} = a_{k1}F_{ij1} + \dots + a_{kR}F_{ijR} + d_k y_{ijk} + \varepsilon_{ijk} \quad (21)$$

where

- R = number of factors common to all items,
- x_{ijk} = person i 's rating of product j on attribute k ,
- a_{k1} = effect of common factor 1 on attribute k (called a loading),
- F_{ij1} = person i 's score of product j on factor 1,
- d_k = weight of unique factor y ,
- y_{ijk} = unique factor of product j on item k for person i ,
- ε_{ijk} = error term.

Thus, in common-factor analysis, the perceptual model has each observed variable being described in terms of a set of R ($R < k$) common factors plus a factor unique to the original observed variable. Generally, the original items are standardized so that certain relationships hold:

The loadings $\{a_{kr}\}$ represent the correlation, ρ , between (hypothetical) factor r and the variable k , and a_{kr}^2 represents the fraction of variance in variable k accounted for by factor r :

$$\rho(F_r, x_k) = a_{kr}. \quad (22)$$

Table 2.4
Perceptual mapping models: extensions and applications

Issues addressed	Model	Data	Illustrative literature	Comments/application
Attribute perceptions: Summarizing perceptions	Factor analysis: Principal components	Attribute perceptions	Hauser & Shugan [1980]	Telecommunications services
Modeling error in perceptual measurement	Probabilistic interpretation of model	Attribute perceptions and preference probability	Bechtel [1985]	Soft drinks
Testing hypothesized relations between variables	Confirmatory factor analysis	Multiple measures of factors	Bagozzi [1977]	
Dissimilarities: Product comparisons	Multidimensional scaling: Single-mode MDS	Similarity judgments	Lehmann [1971]	Soft drinks
Product and preference comparisons	Two-mode MDS	Preference judgments	Green & Rao [1972]	Breakfast choices
Uncertain judgments	Probabilistic interpretation	Preference, degree of preference of concepts	MacKay & Zinnes [1986]	Housing market
Asymmetric brand effects	Asymmetric MDS	Switching data	Harshman, Green, Wind & Lundy [1982]	Car purchase
Alternative data sources:	Logit choice	Choice data	Elrod [1988]	Coffee panel data
			DeSarbo & Hoffman [1987]	Telecommunication equipment
			Katahira [1990]	Car market

Understanding positioning and its effect on share	Regression sales response MDS in time series analysis	Sales panel data and marketing mix	Moore & Winer [1987]
Other attribute-based perceptual mapping methods:			
Analyzing the categories in terms of positioning	Correspondence analysis	Categorical data	Hoffman & Franke [1986]
Understanding the class products belong to	Multiple discriminant analysis	Attribute data and categorical classification	Carroll & Green [1988] Albaum & Hawkins [1983]
Overview and comparison:			
Comparison of MDS approaches			Cooper [1983]
Comparison of factor analysis approaches			Mukherjee [1973]
Comparison of factor analysis and MDS			Hauser & Koppelman [1979] Huber & Holbrook [1979]
			Soft drinks Car purchase Geographic mobility

The *communality* h_k^2 expresses the percentage of the variance in variable k accounted for by the R common factors:

$$h_k^2 = \sum_r a_{kr}^2. \quad (23)$$

The *eigenvalue* λ_r represents the contribution of each factor to the total variance in the original variables:

$$\lambda_r = \sum_k a_{kr}^2. \quad (24)$$

In a specific application it is not uncommon to extract a small number of factors that account for the major part of the total variance.

Another useful aspect of factor analysis is the construction of a perceptual map – the matrix of factor scores – that describes the factor scores as a linear function of the original ratings:

$$F_{ijr} = b_{r1}x_{ij1} + \dots + b_{rK}x_{ijK} + \text{error}, \quad r = 1, \dots, R \text{ for each individual } i. \quad (25)$$

The perceived position of product j is usually constructed by averaging the F_{ijr} over the respondents, i .

An alternative form to common-factor analysis is principal-components factor analysis. Principal-components factor analysis is the same as that expressed in Equation (21) with the exception that the unique factors, y_{ijk} , are omitted. All the variation between the ratings of stimuli are attributed to the underlying factors (F_{ijr}). Studies comparing principal-components and common-factor analysis generally find similar results. For an example of the application of factor analysis in marketing see Hauser & Shugan [1980].

Both the number and the names of factors are important issues in performing a factor analysis. The number of factors used is often chosen based on the magnitude of the eigenvalue of the last factor chosen and the interpretability of the solution. An examination of factor loadings, supplemented by market knowledge, generally leads to reasonable names or interpretations for factors.

5.2. Perceptual mapping: extensions and applications

Table 2.4 outlines some key extensions and applications of perceptual mapping.

The form of common-factor analysis we discussed above is called exploratory factor analysis. It places no constraints on which variables load on the various factors and assumes that all unique factors are independent. In *confirmatory factor analysis* the researcher imposes constraints motivated by theory as to which common factors are correlated, which observed variables affect which factors, and which pairs of unique factors are correlated.

Thus, if we represent the relationship between the variables that we observe (x 's) and the underlying common factors (ξ 's) and unique factors (δ 's) in Figure 2.1, then the relationship between ξ_1 and ξ_2 would not have been permitted in exploratory factor analysis and all the relationships between the factors (ξ) and observed variables (x) would have had to be included.

In confirmatory factor analysis we can constrain some of the a_{kr} 's in Equation (21) to zero or some other value and test that assumption; we can also allow the unique factors $\{\delta_i\}$ to be correlated. If we assume that the common and unique factors are normally distributed, then we can test hypotheses concerning the appropriateness of alternative structures. Confirmatory factor analysis provides a unique maximum-likelihood estimate of a predetermined structure and also provides a chi-square statistic to test the number of factors necessary to account for the correlation matrix. The seminal work in this area was developed by Jöreskog & Sörbom [1979]. An excellent source for the fundamentals of the approach is Long [1983].

A number of studies have compared the application and performance of multidimensional scaling and factor-analytic methods of building perceptual maps. Shocker & Srinivasan [1979] provide a detailed review of the use of these techniques in new product development and concept evaluation. In comparing the performance of factor analysis with multidimensional scaling, Hauser & Koppelman [1979] conclude that factor analysis is superior from the standpoints of predictive ability, interpretability, and ease of use. However, the difference in nature between the two techniques will lead to differences in the prevalence of their application. In markets that are relatively new and in which the cognitive structure of consumers is not well understood or well developed, multidimensional scaling on dissimilarity data might be preferred because it makes fewer assumptions about the criteria on

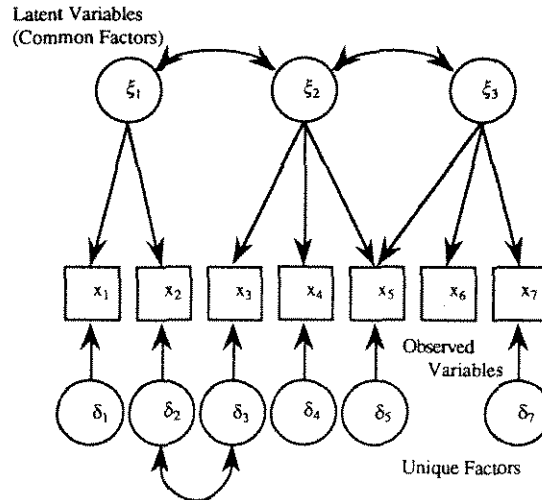


Fig. 2.1. Example of a confirmatory factor analysis model. (Source: Long [1983, p. 14].)

which consumers will evaluate products. Where there is a solid, historical structure for the product category, factor analysis may be preferred since it will add diagnostic richness on which attributes are causing the positioning of products on the perceptual map. In the absence of a well-developed theory of the formation of evaluation criteria, the researcher might well be advised to perform both analyses in search of a convergent picture of the market.

Several other methods have been applied to modeling perceptual spaces, particularly multiple discriminant analysis and correspondence analysis. Multiple discriminant analysis (MDA) is a method of determining which variables explain the groups to which different stimuli belong [Albaum & Hawkins, 1983]. Correspondence analysis is a method that summarizes both the rows and columns of categorical data in a lower-dimensional space (e.g. what are the underlying types of brands in the category and what are the factor dimensions underlying product attributes?). The flexibility of data format is achieved at a cost of not being able to interpret interpoint distances [see Hoffman & Franke, 1986].

5.3. Preference-formation basics

The previous section dealt with *perceptions* – what we believe about products. Here we deal with attitudes and preferences – how favorably disposed we feel toward those same products.

Most models of preference formation transform or map consumer judgments on product attributes to a scalar attitude or preference measure.

Basic models of attitude formation are either *compensatory* or *non-compensatory*. In a compensatory model, the weakness of a brand or product on one dimension can be compensated for by strength on another, and those strengths or weaknesses are combined to determine an attitude toward the brand. In non-compensatory models, usually only a small number of attributes (two or three, say) are used to evaluate a brand, and shortcomings on one attribute cannot be overcome by favorable levels of another.

A basic compensatory model is Bass & Talarzyk's [1972] belief-importance model, building on Fishbein's [1963] theory of attitude formation. In the belief-importance model, the attitude toward the brand is a function of the beliefs about the attributes possessed by the brand weighted by the importance of each attribute:

$$A_o = \sum_i b_{oi} I_i \quad (26)$$

where

A_o = attitude toward any psychological object o ,
 b_{oi} = belief (subjective likelihood) that object o possesses attribute i ,
 I_i = importance of attribute i .

A basic non-compensatory model is the *conjunctive model* where a consumer considers a brand only if it meets certain minimum, acceptable standards on all

of a number of key dimensions. If any one attribute is deficient, the product is eliminated from consideration.

Let

$$\begin{aligned}
 y_{jk} &= \text{perceived level of (key) attribute } k \text{ in brand } j, \\
 T_{jk} &= \text{minimum threshold level that is acceptable (negatively valued} \\
 &\quad \text{attributes such as price that have a maximum level can be multiplied} \\
 &\quad \text{by } -1), \\
 \delta_{jk} &= \begin{cases} 1 & \text{if brand } j \text{ is acceptable on attribute } k, \\ 0 & \text{otherwise,} \end{cases} \\
 A_j &= \begin{cases} 1 & \text{if it is a preferred brand overall,} \\ 0 & \text{otherwise.} \end{cases}
 \end{aligned}$$

Under the conjunctive model we have:

$$\delta_{jk} = \begin{cases} 1 & \text{if } y_{jk} \geq T_{jk}, \\ 0 & \text{if } y_{jk} < T_{jk}, \end{cases} \quad \text{and} \quad A_j = \prod_k \delta_{jk}. \quad (27)$$

Thus, A_j will be non-zero if and only if $y_{jk} \geq T_{jk}$ for all (key) attributes.

5.4. Product preference models: extensions and applications

Extensions of preference models include compensatory models, structural equation models, utility-theory-based models and non-compensatory models (Table 2.5).

5.4.1. Compensatory models

A variety of methods have been advanced for imputing the relative importances of attributes by relating brand preferences to the amount of each attribute that these preferred brands contained. These methods include multiple regression, linear programming, and monotonic analysis of variance. For a review of data-collection techniques and associated estimation methods see Shocker & Srinivasan [1979], Green & Srinivasan [1978], or Horsky & Rao [1984].

Fishbein and others reassessed his original 1963 model to make it more relevant to marketing [Fishbein & Ajzen, 1975], as the extended Fishbein model. The most widely known extension has others, apart from the person making the purchase, influencing the decision in some decision circumstances. In particular,

$$BI = \sum_i a_i b_i + \sum_j SNB_j MC_j \quad (28)$$

where

$$\begin{aligned}
 BI &= \text{behavioral intent,} \\
 SNB_j &= \text{social normative belief, which relates what an individual considers} \\
 &\quad \text{is expected of him or her by an external social group on scale } j,
 \end{aligned}$$

Table 2.5
Preference models: extensions and applications

Issues addressed	Model	Data	Illustrative literature	Comments/application
Preference as a compensatory function of perceptions:	Compensatory models:			
Preference in terms of expected outcomes	Behavioral intention	Intention, outcome values Likelihood of outcomes	Burnkrant & Page [1982]	Blood donations
Incorporating the effect of others' beliefs	Extended Fishbein	As above plus social normative beliefs	Bearden & Etzel [1982]	Various products
Decreasing utility returns to scale	Ideal-point model	Preference and attribute perceptions	Lehmann [1971]	TV viewing
Theoretical relationships between constructs	Structural equation modeling	Attributes, attitude, behavior	Bagozzi [1982]	Blood donations
Non-additive utilities and risk-aversion	Non-linear utility models: Utility theory and decision analysis	Utility of attribute mixes	Keeney & Raiffa [1976] Farquhar [1984] Hauser & Urban [1979]	Health Plans
Non-compensatory preference functions:				
Series of necessary criteria	Conjunctive	Attribute data	Gensch [1987a]	Fertilizer for farmers
Hierarchy of importance of attributes	Lexicographic	Acceptability or preference	Gensch [1987b]	Industrial insulators
Multiphase processing	Phased decision rules		Wright & Barbour [1977]	
Overview/comparison:				
Comparison of compensatory and non-compensatory rules			Johnson & Meyer [1984] Johnson, Meyer & Ghosh [1989]	Apartment choice

MC_j = motivation to comply with these expectations,
 a_i = evaluation (goodness or badness) of attribute i ,
 b_i = belief that object possesses attribute i .

The extended Fishbein model has generally been shown to perform better than the original model, particularly for goods that were publicly consumed rather than privately, and for goods that were more luxuries than necessities [Bearden & Etzel, 1982]. In addition, Wilson, Matthews & Harvey [1975] and others found that attitudes toward the purchase of a brand and behavioral intention were more closely related to behavior than were attitudes toward the brand itself. (Thus, it may be more relevant to ask consumers whether their teeth will get white if they use Ultra Brite than to ask whether they think Ultra Brite whitens teeth.)

An offshoot of the belief/importance model, the ideal-point model, requires a consumer's rating of an ideal brand along with his or her ratings of the actual brands being analyzed (although ideal levels of attributes can be imputed in the same way that importance weights are estimated).

Example: Lehmann [1971] models television-show choice as follows:

$$A_o = \sum_i V_i |B_{io} - I_i|^k \quad (29)$$

where

A_o = overall attitude (preference for a TV show),
 V_i = weight attached to TV-show characteristic i (action, suspense, humor, etc.),
 B_{io} = belief about show on dimension i ,
 I_i = ideal position on dimension i ,
 k = distance metric.

This model is substantially better at predicting behavior than models based on demographic variables. Attributes should be included in an ideal-point form if, beyond some level (the ideal point), there are negative utility returns for further increases in the attribute. With the width of a car, for example, any rational consumer will find some widths too wide and others not wide enough, implying an ideal level between those extremes. With miles per gallon, however, for most consumers it is likely to be the more the better if all other attributes stay the same, so a belief-importance (or vector) formulation would be more appropriate.

5.4.2. Structural modeling of preferences

Factor analysis can be used in a confirmatory way, as well as to explore the relationship between different measures of a number of variables and their underlying constructs. This framework can be extended to test the relationships between the resultant structures. In the models we have seen in this section so far, there is one measure of attitude or preference and we have attempted to understand it in terms of underlying product attributes. In structural equation modeling there are

several physical and/or psychological states and we test the relationship between them and a number of external factors [Bagozzi, 1980].

Thus, a structural equation model may be:

$$y_{ij} = \alpha_i + \sum_l \beta_{il} y_{lj} + \sum_k \gamma_{ik} x_{kj} \quad (30)$$

where

$$\begin{aligned} y_{ij} &= \text{consumer's response on construct } i \text{ for product } j, \\ x_{kj} &= \text{level of attribute } k \text{ for product } j, \\ \alpha_i, \beta_{il} \text{ and } \gamma_{ik} &= \text{parameters.} \end{aligned}$$

5.4.3. Incorporating uncertainty

The models of this section have assumed that we have multiple criteria of varying importance and that we wish to combine them into an overall value or preference function. In the literature on decision theory all these functions would be called 'value functions'. They map a set of attributes, known with certainty, into a function called 'value'. But there are dangers with this simplified approach. What about a new product? Is it reasonable to use the same model to predict choice when the attributes of some products are known with more certainty than those of others?

The specific form of the value function in Equation (26) assumes constant variations in value for each given level of change in the level of a specific attribute. It also assumes that the value of a change in the level of an attribute is independent of the levels of the other attributes. We may require more general value functions if these assumptions are violated. Utility theory and the direct assessment of a utility function across attributes are useful both for testing these assumptions and for developing appropriate models when the assumptions are violated. Utility theory is also useful for understanding how consumers adjust their valuation of alternatives in the presence of uncertainty [see Keeney & Raiffa, 1976].

Hauser & Urban [1979] report an application of utility theory in marketing. They present potential customers for health plans with choices between a moderately attractive plan in which the level of the attributes is known with certainty and an alternative plan in which one of the attributes could be very good or very poor. They estimate consumers' attitudes toward uncertainty by asking consumers how high the probability of the attribute being 'very good' would have to be for the consumers to choose the uncertain plan. For other applications of utility theory see Currim & Sarin [1984] and Eliashberg [1980]. Eliashberg & Hauser [1985] provide an error theory that allows hypothesis testing for utility models under certain distributional assumptions.

5.4.4. Other non-compensatory models

In a *disjunctive model*, instead of preferred brands having to satisfy *all* of a number of criteria (as in a conjunctive model), they have to satisfy any one of a number of criteria. Under the conjunctive model, the consumer may insist on a computer that has lots of memory *and* software. Under the disjunctive model the

consumer may settle for a computer with either a lot of memory *or* a lot of software [Choffray & Lilien, 1980].

Mathematically we can express this as:

$$A_j = \min\left(\sum_k \delta_{jk}, 1\right) \quad (31)$$

where A_j and δ_{jk} are defined as in Equation (27).

It may well be that neither the conjunctive nor the disjunctive rules will give a single preferred brand; either may yield zero or more than one brand, leading to a need for further rules. A *lexicographic model* assumes that all attributes are used, but in a stepwise manner. Brands are evaluated on the most important attribute first; then a second attribute is used only if there are ties; and so forth. Mathematically, if we assume that the attributes are arranged in order from most important to least important, then brand j is preferred to brand m if:

$$y_{1j} > y_{1m} \quad (32)$$

or

$$y_{ij} = y_{im} \quad \text{for } i = 1, \dots, I \quad \text{and} \quad y_{I+1,j} > y_{I+1,m}$$

for $I < \text{number of attributes}$.

A number of other non-compensatory models have been used: see Bettman [1979] for a more detailed discussion of these and other models.

In general, non-compensatory models require individuals to process information by attribute across brands, while compensatory models require consumers to process information within brand across attributes. Since evaluations are simpler and faster in non-compensatory models, it is likely that they are better representations of decision processes for low-involvement goods or for the screening phase when there are many brands, while compensatory models more accurately describe brand evaluations for high-involvement products in more complex decision-making settings (see Bettman [1979] for a review of the supporting empirical literature). A third alternative is where both types of rules are used in sequence, called a *phased-decision rule* [Wright & Barbour, 1977].

6. Purchase

We do not assume that the consumer will always purchase his or her most preferred brand because of measurement error as well as because of variables such as coupons and deals that intervene between the time of measurement of preferences and purchase. Thus, models of purchase have been developed to relate product preferences to purchase probabilities.

6.1. Purchase model basics

In Section 3 on need arousal, we reviewed binary choice models. When there are three or more possibilities, an additional complexity emerges: how does the choice set affect the probability of choice? A seemingly reasonable assumption, referred to as *Luce's axiom* or the *independence of irrelevant alternatives* (IIA) states that the ratio of choice probabilities of any two products does not change when the consideration set changes as long as both of those products are in that set.

Formally, let

$\Pr(a|C)$ = probability of choosing product a when the choice set is C ,
 C' = some set of products that includes C as a subset.

Further, let both a and b be products in consideration set C (and, hence, C'). Then the IIA assumption states:

$$\frac{\Pr(a|C)}{\Pr(b|C)} = \frac{\Pr(a|C')}{\Pr(b|C')} \quad (33)$$

(assuming all denominators are non-zero and that the probability measures are all developed similarly).

Luce [1959] proves that if the above choice axiom holds and if a utility measure, X , exists that is strictly proportional to the choice probabilities (i.e. $X(a, b) = \Pr(a|C)/\Pr(b|C) = X(a)/X(b)$), then

$$\Pr(a|C) = \frac{X(a)}{\sum_{j \in C} X(j)}. \quad (34)$$

If $X(j) = \exp(bV_j)$, then Equation (34) reduces to the multinomial logit model, the multivariate extension of Equation (7a). Alternatively, the multinomial logit model can be derived directly from a set of assumptions analogous to those for the binary logit model, with the key assumption being that the error terms have mutually independent, identical extreme-value distributions.

The IIA assumption is violated when some products being considered are more similar than others. For example, if a consumer is thirsty, he may choose to have a beer or a soft drink. Considering an additional soft drink is likely to have little impact on the 'beer vs soft drink' decision, but will have a major impact on which soft drink, conditional on a soft drink being selected. We will see that violations of the IIA assumption have been handled by grouping products into similar categories (hierarchical models), grouping consumers into homogeneous clusters (segmentation), and modeling departures from IIA explicitly.

Hierarchical models of choice have been developed to address a sequential type of choice process. To illustrate, consider a consumer's choice process for deodorants, using the decision hierarchy illustrated in Figure 2.2. The consumer

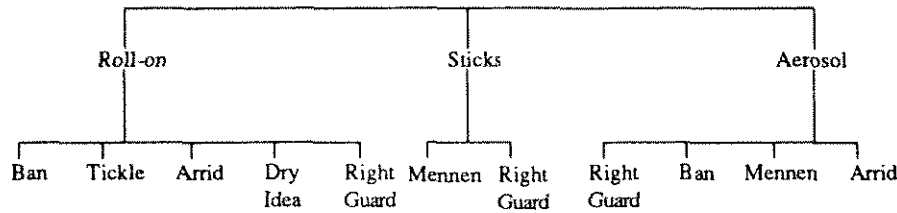


Fig. 2.2. Consumer decision hierarchy for deodorant purchase. (Source: Urban & Hauser [1980, p. 92].)

chooses the form of deodorant and then, conditional on that choice, selects a specific brand. We model the choice of product form and the choice of brand within form separately.

A commonly applied hierarchical model is the *nested logit model*. Algebraically we may write:

$$P_{ij} = P_{ji}P_i \quad (35)$$

where

- P_{ij} = probability of choosing brand j and product form i ,
- P_i = unconditional probability of choosing product form i ,
- P_{ji} = probability of choosing brand j , given product form i .

We assume that utility is separable, as follows:

$$X_{ij} = X_i + X_{ji} \quad (36)$$

where

- X_{ij} = utility from choosing product form i and brand j ,
- X_i = utility associated with product form i ,
- X_{ji} = unique utility of brand j (in product form i).

Brand choice, the bottom level of the hierarchy in Figure 2.2, can be estimated with a multinomial logit model, as before, because it is assumed to satisfy the IIA axiom:

$$P_{ji} = e^{X_{ji}} / \left\{ \sum_k e^{X_{ki}} \right\} \quad (37)$$

The product-form decision may also be modeled using a logit model as long as the error is double exponential and independent of the error at the brand-choice stage. To gain some intuition about the equation for the probability of product-

form purchase, P_i , we note:

$$\begin{aligned} P_i &= \Pr(\max_j X_{ij} > \max_j X_{i'j} \text{ for all } i') \\ &= \Pr(X_i + \max_j X_{ji} > X_{i'} + \max_j X_{ji'} \text{ for all } i'). \end{aligned} \quad (38)$$

From Equation (8) we know that $E(\max_j X_{ji}) = \ln(\sum_j e^{X_{ji}})$. This is called the *inclusive value* of the brand decision in the product-form decision. The equation for the product-form probabilities is given by:

$$P_i = \exp \left\{ \mu \left\{ X_i + \ln \left(\sum_j \exp X_{ji} \right) \right\} \right\} / \sum_{i'} \exp \left\{ \mu \left\{ X_{i'} + \ln \left(\sum_j \exp X_{ji'} \right) \right\} \right\} \quad (39)$$

where μ is normalizing constant.

Note that the individual brand utilities also affect the decision at the product-form level through the inclusive value. Substituting (37) and (39) in (35) yields the nested logit model.

6.2. Purchase models: extensions and applications

Most extensions to these choice-modeling basics have addressed the issues of estimation, of solving the IIA problem and have considered other alternatives for the error term (Figure 2.6).

The important weights in logit models can be derived using a variety of different estimation algorithms. In a review of estimation techniques, Bunch & Batsell [1989] advocate the use of maximum-likelihood procedures. The weights are often called revealed importances because they are revealed by an analysis of choice behavior rather than from direct measurement. They are interpreted in much the same way as regression coefficients. In most computer packages the statistical significance of each importance weight is determined through a *t*-test based on asymptotic values of the standard errors of the estimates. Chapman & Staelin [1982] suggest a procedure that exploits the information content of the complete rank-order choice set.

For marketing applications of the logit model, see Gensch & Recker [1979], Punj & Staelin [1978] and Guadagni & Little [1983]. For more detailed discussion, see McFadden [1976, 1980, 1991].

Louviere & Hensher [1983] and Louviere & Woodworth [1983] provide examples of how the logit model can be combined with experimental design to evaluate hypothetical new product concepts and establish the importance weights of product attributes.

Most applications of the logit model pool data from respondents rather than estimating the model separately for each, in order to obtain sufficient degrees of freedom for the estimation process. Any heterogeneity of consumer tastes can add to the IIA problem because individual customers will have similar brands in their

Table 2.6
Purchase models: extensions and applications

Issues addressed	Model	Data	Illustrative literature	Comments/application
Choice probability in terms of brand utility:	Luce choice model	Utilities and probability of choice	Green, Goldberg & Montemayor [1981]	Career choice
A choice transformation based on specific error theories	Logit choice model	Attribute perceptions and probability of choice	Gensch & Recker [1979]	Store choice
Evaluating new product concepts using experimental designs	Logit choice model	Choice from sets of concept descriptions	Guadagni & Little [1983] Louviere & Woodworth [1983]	Coffee brands Various product categories
Solutions to the independence of irrelevant alternatives problem:				
(a) Creating homogeneous product sets (hierarchical choice models)	Nested logit	Attribute perceptions Choice, product class structure	Dubin [1986]	Heating
(b) Creating homogeneous consumer sets	Segmentation models	Attribute perceptions Choice, knowledge levels	Gensch [1987a]	Fertilizers
(c) Explicitly modeling product interdependence:				
Multiplicative competitive interaction	Luce model of product interactions	Attributes and product choice	Cooper & Nakanishi [1983]	Various products
Logit modeling of product interactions	Generalized logit	Choice from constrained sets Attribute perceptions choice, product-class structure	Batsell & Polking [1985] Dalal & Klein [1988]	Snacks
Normal error theory	Probit	Attributes, choice, perceptual similarity Preference ranking	Currim [1982] Kamakura & Shivastava [1986]	Commuter transport Distribution of ideal points
Overview and reviews			Corstjens & Gautschi [1983] McFadden [1986]	

consideration sets. One way to overcome this source of violation of IIA is to segment the population into homogeneous groups and then estimate a separate model for each segment. Gensch [1984, 1985, 1987a] calls these models estimated at the segment level *disaggregate choice models*. He shows that a priori segmentation, in one case on the basis of knowledgeability and in another on the basis of strength of preferences, improved forecasts and yielded a much richer set of management implications.

Bechtel [1990] shows how the nested multinomial logit model can be estimated when individual-level choice data are not available, but only market-share data in each period. For a more comprehensive development of the nested logit see Ben-Akiva & Lerman [1985]. For applications in marketing, see Dubin [1986] or the Guadagni & Little [1987] example later in this chapter.

In the nested logit model, the decision of whether to buy a specific alternative is made once, at the bottom of the decision hierarchy. In contrast, there are a number of other hierarchical decision models in which the brand is evaluated a number of times on different criteria. Brands in the consideration or choice set are successively discarded until a final choice is made. Models of this type include Tversky's [1972] elimination by aspects model, Hauser's [1986] agenda theory, and Manrai & Sinha's [1989] elimination by cutoffs. These models, while hierarchical choice models, concentrate more on the attribute processing method by which choice occurs than on the structure of the market.

Ensuring that populations are homogeneous and ensuring that products are relatively homogeneous can alleviate problems that the IIA property of the logit model introduces. A further method of addressing those problems is to explicitly model brand interactions. For example, Batsell & Polking [1985] empirically estimate these interactions. Starting with doubles of products $\{i, j\}$ they move on to examine choice probabilities of triples $\{i, j, k\}$, quadruples $\{i, j, k, l\}$, etc. If the ratio of the probability of choosing brand i to brand j is not the same in the triple $\{i, j, k\}$ as it is in the double $\{i, j\}$ then IIA is violated for this triple and a first-order interaction term is included in the model. Similarly, if the ratio of the probability of choosing brand i to brand j is not the same in the quadruple $\{i, j, k, l\}$ as it is in the double $\{i, j\}$, after allowing for the first-order interaction, then a second-order interaction is included. This process continues until including further interactions does not significantly improve fit. Batsell & Polking suggest that this will usually be achieved reasonably parsimoniously (with only first- or second-order interactions being required).

Ben-Akiva & Lerman [1985, p.127] warn that arbitrarily adding interactions can lead to counter-intuitive elasticities. A more axiomatic approach to modeling brand interactions is the generalized extreme-value model developed by McFadden [1978]. For an application of the generalized extreme-value to overcome the IIA problem see Dalal & Klein [1988].

An alternative model to incorporate departures from IIA is the multinomial probit model. This model is an extension of the binary probit developed earlier. It uses a normally distributed error structure and allows the covariance between error terms to be non-zero. But it is not possible to write a general analytical

expression for the choice probabilities, and estimation and evaluation are quite complex. However, recent developments have led to practical computer programs for this model [Daganzo, 1979] and its consequent application in marketing [Currim, 1982; Kamakura & Srivastava, 1984, 1986; and Papatla & Krishnamurthi, 1991].

7. Post-purchase attitudes and behavior

Consumer behavior is an ongoing process: how a brand performed relative to the consumer's needs and expectations triggers what that consumer is likely to do on future purchase occasions. In addition, especially for durable goods, a consumer is likely to communicate his or her level of satisfaction to others, influencing the behavior of future consumers.

7.1. Post-purchase analysis basics

The modeling of consumer satisfaction is based on the confirmation/disconfirmation paradigm. Confirmation occurs when the consumer's perception of how the product performs after purchase matches the expectation the consumer had prior to purchase. Positive disconfirmation occurs when product performance exceeds expectations; negative disconfirmations occur when the product falls below expectations. Satisfaction goes down as the level of negative disconfirmation goes up. Thus,

$$S_t = f(D_t) = g(E_{t-1}, P_t) \quad (40)$$

where

$$\begin{aligned} S_t &= \text{satisfaction at time } t, \\ D_t &= \text{disconfirmation at time } t, \\ E_{t-1} &= \text{expectation prior to experience at } t-1, \\ P_t &= \text{perceived performance (post-experience) at time } t. \end{aligned}$$

Swan & Combs [1976] and Oliver [1980] pioneered the disconfirmation idea. Swan & Combs suggested that disconfirmation could be of two types: instrumental (based on how the product performs functionally) and expressive (related to the feelings associated with the consumption experience).

Howard & Sheth [1969] suggested that the effect of disconfirmation of post-purchase attitudes can be written as:

$$A_{t+1} = h(P_t - A_t) + A_t$$

where A_t is the attitude of the consumer toward the brand. Oliver points out that if A_t is defined in expectational terms such as in a Fishbein model, the equation above provides a mechanism for updating attitudes on the basis of experience. He

notes that two processes are going on here: (1) the formation of expectations and (2) the disconfirmation of those expectations through performance comparisons. Much of the recent literature on customer satisfaction [see Zeithaml, Parasuraman & Berry, 1990, for example] has addressed these two issues in addition to measuring the effect of satisfaction on subsequent behavior.

7.2. *Post-purchase attitudes and behavior: extensions and applications*

The specification and formalization of Equation (40) has led to models of satisfaction, of variety seeking and purchase feedback as well as network/communication models (Table 2.7).

7.2.1. *Satisfaction*

Researchers have primarily related satisfaction to perceived product performance through linear models. For example, Bearden & Teel [1983] used structural equation models (LISREL) to explain how expectations, attitudes, intentions and product experience (disconfirmation) determine satisfaction, along with future attitudes and behavior. They only collected data on perceptions relative to expectations to measure degree of disconfirmation; Churchill & Suprenant [1982] develop similar results with separate measures of perceptions and expectations.

Oliver & DeSarbo [1988] test the effect of attribution, expectation, performance, disconfirmation and equity on satisfaction, using an analysis of variance model. They report interactions between constructs in their analysis and are able to segment their sample on the basis of the decision rules the consumers follow. Tse & Wilton [1988] review the customer satisfaction literature and test different decision rules (stated disconfirmation versus perceived performance minus expected performance) as well as difference benchmarks for comparison (expected performance versus ideal level of performance versus equitable performance).

This area is seeing much attention currently, with researchers focusing on the linearity of the relationships (Woodruff, Cadotte & Jenkins [1983] suggest that a zone of indifference exists), the dynamics of customer satisfaction, and the relationship of customer satisfaction to the purchase experience (via the concept of transaction utility, Thaler [1985]). The literature on service quality is evolving rapidly as well [Parasuraman, Zeithaml & Berry, 1985] and clearly relates closely to the satisfaction construct [Bolton & Drew, 1991].

Word-of-mouth has been studied extensively in sociology [Rogers, 1983] and marketing [Westbrook, 1987]. Biehal [1983] also tests the impact of product experience on search. In the extreme, dissatisfaction leads to complaints to the supplier or to other consumers [Singh, 1988].

7.2.2. *Models of variety seeking*

Stochastic learning models provide one way to relate past purchase patterns to future behavior [see Lilien, Kotler & Moorthy, 1992, Chapter 2]. Stochastic models concentrate on the random element of consumer behavior. In contrast, the

Table 2.7

Post-purchase attitudes and behavior: extensions and applications

Issues addressed	Model	Data	Illustrative literature	Comments/application
Effect of usage on satisfaction and attitude	Structural equation modeling	Expectations Perceived performance Satisfaction	Churchill & Surprenant [1982] Bearden & Teel [1983] Woodruff, Cadotte & Jenkins [1983] Oliver & Desarbo [1988]	Video disk players Auto repair No empirical testing Stock investments
Negative word-of-mouth	Discriminant analysis	As above plus post-purchase behavior As above	Richins [1983] Westbrook [1987] Biehal [1983] Singh [1988]	Clothing/appliances Auto/CATV Auto repair Various categories
Variety seeking and purchase-event feedback: Variety seeking	Logit model Luce model	Purchase over time Attributes and preferences As above	McAlister [1982] Latin & McAlister [1985] Latin [1987]	Soft drinks Soft drinks Soft drinks
Social networks/communication patterns:				
Analysis of network structures	Individual-level diffusion models	Adoption of others plus communication links	Coleman, Katz & Menzel [1957] Czepiel [1974] Burt [1987] Midgeley, Morrison & Roberts [1992] McAlister & Pessemlier [1982] Kahn, Kalwani & Morrison [1986]	Effect of network structures on diffusion of innovations
Review of variety seeking				

variety-seeking literature models the effect of current choice on future behavior by understanding the deterministic influences of choice. McAlister & Pessemier [1982] distinguish between several types of behavior by individuals that relate to multiple needs, the acquisition of information, and the alternating purchase of familiar products (variety seeking). They hypothesize that consumers have an ideal point or satiation level for the product's attributes that leads to its decreasing utility after a period of sustained consumption. Thus, if a man drinks six colas there is a good chance that on his next consumption occasion he might wish for a lemon soda 'just for a change', i.e. perform variety seeking.

Let us consider the study by Lattin & McAlister [1985], who used a Luce model to understand the effect of brand similarities on purchase-event feedback. They model a consumer's utility for a brand on a given consumption occasion as diminishing proportionally to the value of the features it shares with the brand the consumer chose on the previous consumption occasion via a first-order Markov scheme.

Thus,

$$V_{ij} = V_i - \lambda S_{ij} \quad (41)$$

where

$$\begin{aligned} V_{ij} &= \text{utility of } i \text{ given that } j \text{ was chosen previously,} \\ V_i &= \text{unconditional utility of } i, \\ \lambda &= \text{discount factor indicating consumer's variety-seeking intensity,} \\ S_{ij} &= \text{value to the customer of all want-satisfying features shared by } i \\ &\quad \text{and } j. \end{aligned}$$

Applying the Luce model to this formulation gives the probability of purchase of i given a previous purchase of j , P_{ij} as:

$$P_{ij} = (V_i - \lambda S_{ij}) / \sum_k (V_k - \lambda S_{kj}) \quad (42)$$

Lattin & McAlister scale the V_i so that $\sum_k V_k = 1$. Then V_i can be interpreted as the probability of choosing brand i in the absence of variety seeking.

The authors analyze how much previous consumption alters the unconditional brand-choice probability, V_i ; that is, $P_{ij} - V_i$. Now $P_{ij} - V_i < 0$ indicates that product j is a substitute for product i (the consumption of brand j lowers the probability of choosing brand i), while $P_{ij} - V_i > 0$ suggests that brand j is a complement to brand i (the consumption of brand j increases the probability of choosing brand i). An examination of Equation (42) shows that brand j will be a substitute for brand i ($P_{ij} - V_i < 0$) if $V_i < S_{ij} / \sum_k S_{kj}$, that is, if product i shares more than a proportional amount of its want-satisfying value with product j .

The asymmetry of the effect of product i on product j , and product j on brand i , may be seen by expanding the expression for the effect of product j on i ($P_{ij} - V_i$) using Equation (42). Now $P_{ij} - V_i$ may be rewritten:

$$P_{ij} - V_i = \lambda \left(V_i \sum_k S_{kj} - S_{ij} \right) / \left(1 - \lambda \sum_k S_{kj} \right). \quad (43)$$

Since $S_{ij} = S_{ji}$, $P_{j|i} - V_j$ is (by symmetry)

$$P_{j|i} - V_j = \lambda \left(V_j \sum_k S_{ki} - S_{ij} \right) / \left(1 - \lambda \sum_k S_{ki} \right). \quad (44)$$

The intensity of variety seeking λ , will magnify the degree of substitution or complementarity. Lattin & McAlister report the results of analyzing the soft-drink consumption behavior of 27 students over a period of 81 days. They obtain a least squares fit to Equation (43) using constrained non-linear programming, estimating their model at the individual level and averaging across individuals to obtain segment-level estimates. The variety-seeking segment consists of 9 of the 27 respondents with λ 's ranging from 0.43 to 0.96. For the other 18 respondents a simple Markov model was sufficient.

In other models of variety seeking, McAlister [1982] considers multiperiod depletion of inventories of attributes, Lattin [1987] explicitly models different levels of variety seeking for different attributes using a logit framework, and Simonson [1990] examines the contextual determinants of variety seeking.

7.2.3. Models of social communication networks

Most aggregate levels of the diffusion process either implicitly or explicitly assume perfect (or at least random) mixing between the members of the population. With models at the individual level we can investigate this assumption. The pioneering work examining the structure of interpersonal networks through which word-of-mouth about innovations diffuse was the study of the diffusion of prescriptions of the drug tetracycline in four Illinois towns by Coleman, Katz & Menzel [1957]. Cziepel [1974] shows that for innovations in steel production not only did specific incomplete networks exist but there were cliques with good communications within them but weak communications between them. Burt [1987] goes further in suggesting that for many innovations the diffusion effect is not directly from adopter to potential adopter but indirectly through some third party who is not part of the network. Midgley, Morrison & Roberts [1992] develop a model of the effect of network structure on the internal diffusion process of firms and test the degree to which departures from the perfect mixing model occur. In addition, they test the degree to which these departures from traditional assumptions at the individual level affect the diffusion process at the aggregate level. They find evidence of cliqueing and determine that it could cause anomalous aggregate-level diffusion patterns. While all of these studies were conducted in industrial markets because of the difficulty in specifying networks for consumer products, the theory is also likely to be valid for consumer markets.

8. Integration: Combining models to solve management problems

We have now examined some of the major tools to understand and predict consumer behavior using process-oriented models. In developing our framework we emphasized the need for parsimony and ease of implementation. Thus, we

looked at different stages of the decision process separately to concentrate on only those phenomena that are keys to understanding behavior and product management. However, there are many situations in which more than one stage and more than one type of model are appropriate.

In this section we give three illustrations of how the building blocks that we have developed can be joined and applied to address particular real-world consumer purchase situations. Our illustrations include: (1) the combination of purchase incidence (need arousal) and brand choice (purchase) models, (2) joint consideration and purchase models, and (3) models of information integration and evaluation.

8.1. Application 1: A model of brand choice and purchase incidence

Guadagni & Little [1987] examine how a retail store should price, display and discount products to maximize sales and contribution. To achieve that end, the management needs to understand the source of additional product sales: do those sales come from increases in brand share or from additional sales of the category as a whole? The areas of our process model that correspond to these issues deal with the purchase stage (the brand that will be chosen) and the need-arousal stage (when a category purchase will occur).

Both phenomena can be modeled using logit models: earlier we developed the nested logit model which provides an integrated framework to understand related logit decision processes. Guadagni & Little use this technique to study purchase incidence and brand choice. They examine the purchase of coffee in Kansas City over a 74-week period using scanner panel data on a sample of 200 households.

Their model of brand choice follows their earlier work [Guadagni & Little, 1983]. First, they identify the determinants of *brand choice* for the eight major brand-size combinations of coffee (products) on the market:

- X_{1ijt} = brand loyalty in period t of customer i to the brand-size of product j ,
- X_{2ijt} = size loyalty in period t of customer i to the size of brand-size j ,
- X_{3jt} = the presence of a promotion (display) on brand-size j in time t ,
- X_{4jt} = the discount of brand-size j at time t as a proportion of the average category price,
- X_{5jt} = the regular (undiscounted) price of brand-size j at time t relative to the average category price.

They also develop seven product-specific dummy variables to capture the brand equity of each of the products.

The logit brand-choice model estimated across respondents, products and time periods, is:

$$P_{ij|B} = \exp \left\{ \sum_n \beta_n X_{nijt} + \sum_m \gamma_m \delta_{imt} \right\} / \sum_k \exp \left\{ \sum_n \beta_n X_{nikt} + \sum_m \gamma_m \delta_{ikt} \right\} \quad (45)$$

where

- $P_{ijt|B}$ = the probability of individual i purchasing product j in time period t given a purchase in category B,
 β_n, γ_m = coefficients of the independent variables and product-specific dummies respectively ($n = 1$ to 5 and $m = 1$ to 7).

Their model of purchase incidence is a binary logit, with the dependent variable being whether a purchase took place or not. Their explanatory variables are:

- Z_{1Bi} = a dummy variable for the utility that consumer i gets from buying (B) in the category,
 Z_{2Bit} = a variable to denote whether consumer i made multiple purchases when buying on shopping trip t ,
 Z_{3Bit} = household inventory of coffee,
 Z_{4Bit} = the category attractiveness,
 Z_{5Bit} = the average category price,
 Z_{6Bit} = a dummy variable to account for an announcement of impending price rises due to a crop failure in Brazil.

Household inventory, Z_{3Bit} , can be estimated from purchase history and average seasonal consumption rates. Category attractiveness, Z_{4Bit} , is the inclusive value from the nested logit:

$$Z_{4Bit} = \ln \left\{ \sum_k \exp \left\{ \sum_n \beta_n X_{nikt} + \sum_m \gamma_m \delta_{imt} \right\} \right\} \quad (46)$$

To calculate the expected category attractiveness of buying later (of *not* buying in this time period), Z_{4Nit} , and the expected price that consumer i would face if he or she bought later (and *not* this period), Z_{5Nit} , Guadagni & Little take an average of the attractiveness and category prices over the previous eight purchase occasions, respectively. For all other variables, ($Z_{1Nit}, Z_{2Nit}, Z_{3Nit}, Z_{6Nit}$), they set the utility of not buying in period t to zero.

Their final binary logit model of whether category purchase will occur on shopping occasion t or not is:

$$P_{Bit} = \exp \left\{ \sum_n \alpha_n Z_{nBit} \right\} / \left\{ \exp \left\{ \sum_n \alpha_n Z_{nBit} \right\} + \exp \left\{ \sum_n \alpha_n Z_{nNit} \right\} \right\} \quad (47)$$

where P_{Bit} is the probability of buying (B) within the category on shopping occasion t and α_n ($n = 1, 2, \dots, 6$) are parameters.

Guadagni & Little estimate the model by first calibrating Equation (45), the brand-choice equation, using maximum-likelihood techniques. The 'inclusive value' can then be calculated from Equation (46). This enables the purchase-incidence binary logit model, Equation (47), to be calibrated.

At the product level (brand choice), both brand-name loyalty and product-size loyalty are strongly statistically significant. The brand-choice decision also responds to store promotions and price cuts, as well as being sensitive to the usual price of the brand relative to the market. At the purchase-incidence level, all variables except the price of the category are statistically significant, suggesting that coffee purchase does not depend on the price of the category as a whole, but that brand choice is sensitive to price.

Their combined brand-choice/purchase-incidence nested logit model allows forecasts both of brand share and of total brand sales. The effects of individual products' marketing activity are traced through to category effects using the 'inclusive value' from the nested logit. Additionally, their purchase-incidence model enables them to evaluate the effect of external factors, such as multiple purchases, household inventory, and crop failure in Brazil.

By approaching the same problem from a stochastic modeling perspective and incorporating explanatory variables into their model, Hauser & Wisniewski [1982] and Wagner & Taudes [1986] jointly model purchase-incidence and brand-choice decisions and their determinants. Hybrid approaches combining discrete choice and stochastic models include those of Jones & Zufryden [1981] (with a negative binomial distribution purchase-incidence model and logit brand choice) and Gupta [1988] (with an Erlang interpurchase time distribution and logit brand choice). Gupta also models purchase quantities in his framework using a cumulative logit model.

8.2. Application 2: Integrating consideration and choice

Gensch [1987b] studies the management problem of influencing the choice process of electric utilities for an industrial durable in order to better design and position the product. He cites considerable work applying compensatory evaluation rules to examine the relation between product positioning, perceptions and purchase. But evidence in the consumer behavior literature suggests that consumers often use a two-stage procedure [e.g. see Wright and Barbour, 1977]. In our framework this concept corresponds to a screening phase in the information search stage to determine the consideration set, followed by an evaluation and/or purchase stage. Gensch proposes a non-compensatory screening phase, followed by a compensatory evaluation of, and choice between, the surviving alternatives.

8.2.1. Attribute-based screening model

In Gensch's models, consumers screen products sequentially by attribute on a conjunctive basis. That is, consumer i ensures that his or her perception of product j on attribute k , y_{ijk} , suitably scaled, does not fall too far short of the best brand's level on that attribute. Starting with the most important attribute, the consumer screens all brands on this basis to come up with an acceptable consideration set. These brands then enter the purchase phase. Mathematically the model may be written as follows.

First the attribute levels, y_{ijk} , are rescaled by

$$x_{ijk} = [\max_n y_{ink} - y_{ijk}] / \max_n y_{ink} \quad (48)$$

where x_{ijk} is the rescaled perception that consumer i has of attribute k for product j . Gensch then postulates that there are maximum levels of x_{ijk} that the consumer will find acceptable; he calls these thresholds T_k . He assumes the utility that attribute k offers is

$$v_{ijk} = \max(0, T_k - x_{ijk}).$$

Gensch then uses a multiplicative utility function to determine brand j 's overall utility to consumer i , V_{ij} :

$$V_{ij} = \prod_k v_{ijk}. \quad (49)$$

Equation (49) implies that if a brand fails to meet the threshold on any one criterion, then it has zero utility, i.e., $V_{ij} = 0$. (A conjunctive decision rule.) He estimates the T_k from the data to maximize:

$$\prod_i \left(V_{ij^*} / \sum_j V_{ij} \right)^{X_i} \cdot \left(1 - V_{ij^*} / \sum_j V_{ij} \right)^{1 - X_i} \quad (50)$$

where

$$j^* = \text{the chosen alternative,}$$

$$X_i = \begin{cases} 1 & \text{if } j^* \text{ is not screened out in the consideration phase,} \\ 0 & \text{otherwise.} \end{cases}$$

Given estimates of thresholds, T_k , the consideration set consists of brands such that $V_{ij} > 0$, i.e. brands that do not fail any of the threshold criteria. If no brands fulfill this criterion, the consumer chooses the brand(s) eliminated last (i.e. that failed the least important conjunctive criterion) [see Gensch and Svestka, 1984].

8.2.2. Logit discrete-choice model

Given a consideration set from the previous stage, choice data, and respondent perceptions of suppliers on the salient attributes, Gensch fits a standard logit choice model. The probability of purchase of brand j by consumer i , P_{ij} , is given by:

$$P_{ij} = \exp \left\{ \sum_k b_k y_{ijk} \right\} / \sum_{n \in C} \exp \left\{ \sum_k b_k y_{ink} \right\} \quad (51)$$

where C is the set of brands surviving the screening process.

8.2.3 Model testing

Gensch tests the model by examining the ratings of four suppliers by 182 buyers of electrical generation equipment on eight attributes, combined with choice data.

The two-stage model gives superior predictions to several a one-stage models. In addition, two attributes dominate the screening process. Of all eliminations 70% occur on the basis of manufacturer quality, while 26% occur on the basis of manufacturer problem-solving ability. The one-stage logit model has both of these variables statistically significant. After incorporating a screening phase, both variables are insignificant. This suggests that these variables are important in gaining consideration, but once consideration is gained, they do not play a significant role in evaluation and choice. The use of this and related models had a major impact on the successful performance of ABB Electric in the 1970s and 1980s [Gensch, Aversa & Moore, 1990].

One of the strengths of Gensch's approach is that it only requires attribute perceptions, perceived importance and choice data, making the data-collection task easier than that for the consideration models described earlier. It also means that care has to be taken in interpreting the screening results: we cannot be sure that non-compensatory screening was in fact the process that took place. Rather, we can only conclude that a non-compensatory screening model combined with a compensatory choice model gives a better fit than does either a single-stage compensatory model or a discrete-choice model.

8.3. Application 3: The dynamics of perceptions, preference and purchase

Roberts & Urban [1988] develop a dynamic brand-choice model to address the problem of forecasting sales of a new consumer durable. With the launch of a new automobile, appeal (perceptions, preference and choice) is important, but so are the dynamics of how that appeal will change as the product diffuses through the adopting population. GM's Buick Division was interested in how word-of-mouth and information from other sources would affect sales. The modeling task combines static choice modeling with a model of information integration to help understand the dynamics of choice.

8.3.1. Probability of purchase model at any point in time

Roberts & Urban use decision analysis theory to show that a (rational) risk-averse consumer with uncertain beliefs will attempt to maximize the expected utility that he or she will get from the brand minus a constant times the uncertainty involved [Keeney & Raiffa, 1976]. They term this quantity the risk-adjusted preference, x . Thus, the risk-adjusted preference for brand j , x_j , is given by:

$$x_j = V_j - (r/2)\sigma_j^2 \quad (52)$$

where

$$\begin{aligned} V_j &= \text{the expected utility from brand } j, \\ \sigma_j^2 &= \text{the variance of beliefs about } V_j, \\ r &= \text{the consumer's risk-aversion.} \end{aligned}$$

V_j may be further modeled in terms of its constituent attributes to incorporate product positioning.

If x_j is measured with error following the extreme-value distribution, we can model the probability of choosing brand j in logit terms:

$$P_j = e^{x_j} / \sum_{k \in C} e^{x_k} \quad (53)$$

By substituting Equation (52) into (53) we can see how perceptions, expected preference and the level of uncertainty affect the probability of purchase.

8.3.2. Dynamics of preference uncertainty and choice probability

The literature on diffusion of innovations suggests that information fills two roles: it reduces uncertainty and it can lead to changes in attribute perceptions and thus preference. To model the effect of new product information on an individual's perceptions and uncertainty, Roberts & Urban use Bayesian updating theory. Beliefs about expected preference for the brand after search are a weighted average of the prior beliefs and the level of preference that the new information suggests. Uncertainty is reduced by an amount that depends on the faith that is placed in new information. Mathematically,

$$V_j'' = (\sigma_{wj}^2 V_j' + \sigma_j'^2 V_{wj}) / (\sigma_{wj}^2 + \sigma_j'^2) \quad (54)$$

and

$$\sigma_j''^2 = \sigma_j'^2 \sigma_{wj}^2 / \{\sigma_{wj}^2 + \sigma_j'^2\} \quad (55)$$

where

- V_j'' and $\sigma_j''^2$ = expected preference and uncertainty associated with product j after updating,
- V_j' and $\sigma_j'^2$ = prior expected preference and uncertainty,
- V_{wj} and σ_{wj}^2 = average preference and uncertainty of the incoming word-of-mouth about product j .

By substituting the dynamics of expected preference and uncertainty (Equations (54) and (55)) into the equation for risk-adjusted preference (Equation (52)) and the probability of purchase (Equation (53)), Roberts & Urban derive a dynamic brand-choice equation that provides a model of the individual-level changes that are driving the aggregate-level diffusion process. They assume that the rate at which new information about product j will become available is proportional to the number of cumulative adopters at time t , Y_{jt} . The variance of word-of-mouth information will be inversely proportional to this rate and thus the cumulative sales of product j :

$$\sigma_{wj}^2 = k_j / Y_{jt} \quad (56)$$

where k_j is a constant reflecting the salience of the new product, j .

8.3.3. Application

In an empirical application, Roberts & Urban calibrated their model using 326 respondents' evaluations of the existing US automobile market and a new product concept. After measuring perceptions, uncertainty, and purchase probabilities with respect to the existing market, respondents were exposed to increasing amounts of information about the new car. First, they saw a description, then they took a prototype for a test drive followed by a videotape of owners' reactions to the new car, and finally they were given a consumer report written by an independent testing service. Measures of perceptions, preference, uncertainty, and probability of choice were taken after every information exposure to calibrate the model dynamics.

Respondents were exposed to either positive- or negative-information videotapes. Roberts & Urban traced the effect of these changes through the risk-adjusted preference to calculate the effect of word-of-mouth on individuals' probability of purchase and thus the diffusion of the new product, using Equations (52) and (53).

The results of the perceptual evaluation stage of the model helped shape the advertising copy for the car, stressing that the perceived durability did not come at the expense of comfort and style. An understanding of the dynamics of expected sales of the car under conditions of negative word-of-mouth persuaded management to delay the launch for over six months until a transmission fault was fixed. Finally, although it is difficult to validate durable forecasting models, preliminary indications are that this methodology predicted actual market performance well [Urban, Hauser & Roberts, 1990].

Meyer & Sathi [1985] propose a similar model with exponential-smoothing updating for beliefs and uncertainty. Recent work is addressing the problem of making these dynamic brand-choice models more parsimonious in parameters to allow the aggregate effects of individual-level changes to be examined more readily [Oren & Schwartz, 1988; Chatterjee & Eliashberg, 1990; Lattin & Roberts, 1988].

9. Future directions for consumer behavior models

The field of modeling consumer behavior is evolving rapidly. Changes are being driven by advances in our theoretical understanding of consumer decision processes (e.g. developments in the use of discrete choice models to overcome the independence of irrelevant alternatives), the availability of new data sources (e.g. the new mapping techniques that take advantage of the existence of longitudinal supermarket scanner data at the individual level), and new or more focused concerns of managers (e.g. models of market structure that address problems of market coverage and portfolio management).

Our study of the consumer decision process above also highlights a number of under-researched areas, particularly in terms of need arousal, information integration and post-purchase influences. As the external environment changes at an increasing rate and product life-cycles become shorter, equilibrium models of the five stages of our process are likely to become less popular than those incorporating dynamics.

We expect advances in consumer behavior models in a number of key areas, and we highlight four.

9.1. Accuracy and applicability

Currently most of our models still take a reasonably narrow view of the utility-maximizing consumer, though this understanding has broadened over the last few years. Our models must handle better the phenomena of simplifying heuristics, biases in decision-making, and the effects of framing and context as we broaden our modeling perspective beyond that of a strictly rational consumer.

Simplifying heuristics have been well recognized in the consumer behavior literature for some time [e.g. Lindblom, 1959]. To some extent, the non-compensatory models that we discussed here represent a beginning in this area. However, there is no satisfactory error theory associated with those rules and also there are many simplifying heuristics that consumers employ that are not covered by those models. The situations that trigger different heuristics have attracted attention in consumer behavior [e.g. Payne, Bettman & Johnson, 1988]. In order to quantify these effects, marketing scientists must formalize them in mathematical terms.

Recent research in the area of psychology into the *biases* that consumers exhibit are likely to modify the strict utility-maximizing models that we often apply. Tversky & Kahneman [1974] observe these biases in decision-makers (including consumers): representativeness (in which perceived ideas and stereotypes obstruct learning of new information), availability (in which respondents select events or outcomes that are familiar to them, even in defiance of the true probabilities), and anchoring (in which respondents are reluctant to adjust adequately from their preconceived ideas). Subsequently, Kahneman & Tversky [1979] have also studied asymmetric risk-aversion to gains and losses from the consumer's current position and the over-weighting of low-probability events. Currim & Sarin [1989] develop a method of calibrating and testing these prospect theory models. They show that for paradoxical choices, prospect theory [Kahneman & Tversky, 1979] outperforms utility theory, while for non-paradoxical choices there is little difference. Winer [1986] and Lattin & Bucklin [1989] include reference levels in their models of brand choice to account for anchoring, but little work has been done to model and develop an error theory for the simplifying heuristics that consumers appear to use to make decisions.

Stemming from the work of Kahneman & Tversky above, there is a wide variety of other contextual effects that influence choice, most of which are currently omitted from our models. Meyer & Eagle [1982] show that contextual factors influence the importance weight that is given to attributes. The range of alternatives available, for example, is important in choice [e.g. Alba & Chattopadhyay, 1985, 1986]. The variety-seeking literature also studies one particular form of context. Other interesting work in context includes Thaler's [1985] mental accounting model. Thaler suggests that there are two forms of utility associated with every consumer decision: acquisition utility and transaction utility. The former is the utility modeled in this chapter. Transaction utility reflects the approval or pleasure that consumers

gain from the transaction itself. Transaction utility is particularly tied to the concept of fairness and appropriateness. While Urbany, Madden & Dickson [1989] find little evidence of fairness affecting behavioral intentions, there are clearly situations in which transaction utility will be a key determinant of choice.

9.2. Modeling the mental process

The field of consumer behavior is advancing our understanding of the mental processes that consumers undertake. We must begin to incorporate the role of memory into our models, together with the effects of categorization on consumer choice. Also we must determine the influence of too much or too little information (information overload and inferencing, respectively).

We predict that formal modeling of mental processes, *memory* in particular, will see some important advances. Lynch & Srull [1982], for example, have demonstrated the importance of the role of memory in consumer behavior experiments. Information once remembered will not necessarily be recalled and recent work looks at the salience of brands and stimuli [e.g. Nedungadi, 1990].

The study of how a product is perceived and how consumers form personal product classes is called *categorization* [e.g. Sujan, 1985]. The product class then becomes the benchmark against which the product is evaluated, either by use of a prototypical product from the category for comparison, or with an exemplar. The implications of how products are categorized extend to the probability of consideration [Urban, Hulland & Weinberg, 1990], managing cannibalization (competition with the firm's other brands), brand draw (competition with competitors' brands), and product-line extensions [Bridges, 1991].

A considerable body of literature exists on how consumers react when there is incomplete information, or when there is too much information: *information overload and inferencing*. Huber & McCann [1982] show that consumer inferences affect choice, and a number of researchers have developed models to explain this [e.g. Malhotra, 1986]. Dick, Chakravati & Biehal [1990] demonstrate that consumers infer from memory-based knowledge when faced by incomplete external information. Malhotra [1982] and Keller & Staelin [1989] show that information overload also affects choice, though less work has been undertaken to model the simplifying heuristics that the consumer uses to overcome the problem.

9.3. Coping with consumer heterogeneity

Many of our individual models are estimated across all respondents and assume perfect homogeneity to obtain sufficient degrees of freedom. Both by grouping similar people together and estimating at the segment level, and also by estimating some parameters within respondents and others across respondents, we can cope with consumer heterogeneity via richer models.

Much of the consumer behavior work to date has postulated a single model of consumer response and tested its fit in a given situation against some competing, single model. We expect to see more work on methods for testing for this assumed

customer homogeneity, applying models to homogeneous segments, i.e. *matching models to market segments*. For example, Gensch [1985] shows that prior segmentation of a population into segments homogeneous in perceptions and preferences gave much higher predictive accuracy than when a single model was fitted to the entire population. Such segmentation is particularly powerful when it can be done a priori, as Gensch [1987a] shows on a population of industrial buyers. In this latter application, more knowledgeable buyers are modeled well within a logit model while less knowledgeable buyers closely approximate the assumptions of a hierarchical choice model. We expect that further research will generate a better understanding of what types of models are most likely to be effective a priori, leading to the better application of appropriate, situation-specific consumer behavior models.

9.4. Markets and consumers in flux

Market dynamics need to be incorporated in a number of ways. We have already spoken of the move from modeling demand using comparative statics to using dynamic models. That move will continue. Additionally, the supply side of market dynamics will attract greater attention, as will decisions involving unfamiliar products.

Supply-side feedback, in the form of the game-theoretic implications of a competitor's actions on an organization's optimal strategy are discussed by Moorthy in Chapter 4. However, in many markets, product design evolves over time, reflecting the interplay between customers' needs and the company's manufacturing capability. While Hauser & Simmie [1981] made a start on modeling this process, there is still much that needs to be done.

The problems of decisions involving *unfamiliar alternatives* (e.g. new, technologically sophisticated products) is not well modeled or measured in marketing at present. There is some evidence that novices do not use as rich a mental model as experts [e.g. Alba & Hutchinson, 1987]. Management needs in high-technology areas will continue to push for the development of better methods to assess how consumer decision processes evolve in such markets.

A modeling challenge when studying unfamiliar products is that the benchmarks that consumers use in these areas are generally not comparable to the unfamiliar product. For example, consumers might compare a compact disk player to a stereo cassette player. Johnson [1984, 1986, 1988] has suggested that the more consumers are faced with non-comparable alternatives, the more they resort to hierarchically based processing. That is, he asserts that with non-comparable alternatives, consumers evaluate brands by a very general comparison, rather than a specific attribute-based one. As markets evolve more rapidly and product life-cycles shorten, the need for models to understand the adoption of truly new products will become more evident.

While there will undoubtedly be many other phenomena that will be studied, leading to a more complete understanding of consumer behavior, we have attempted to sketch what we see as some of the major areas. These advances will

occur in conjunction with a study of the context in which purchase and consumption take place. Thus, dyadic and family behavior will develop as an example of sociological influences on decision-making. Hedonic behavior and the consumption of services will provide an example of product context influencing behavior. And distribution channel and salesforce influences will shape research on physical context affecting behavior. There will be other advances addressing specific marketing problems (such as the marketing of groups of products as a portfolio in the financial services industry), but we feel that improving the accuracy and validity of our models, ensuring their fit to different customers and the stability of that fit over time will bring the greatest rewards to researchers and managers over the next five years.

10. Conclusions

In a field as vast and diffuse as that of consumer behavior and consumer markets, it is difficult to develop a single best synthesis. We have drawn from developments in the literature of the behavioral sciences, economics, marketing, statistics, and the like and have categorized according to the stage or stages in the decision process to which those developments appear most applicable. Our models overlap these processes and stages but this integrating framework provides a useful way of organizing this large, diverse literature.

We followed our framework from need identification to information search to perception and evaluation and preference formation. From there we went on to purchase and post-purchase feedback models. Finally, we demonstrated the power of developing models that combined several of the above stages in a single framework.

The future of consumer behavior modeling is bright; newer models are richer, more flexible, and more closely attuned to modern data sources. Yet many phenomena are poorly modeled at the moment. We highlighted the areas of modeling consumer purchase heuristics (and information-processing biases), modeling consumers' mental processes, matching models to market segments, and modeling choice for truly new or non-comparable alternatives as fruitful areas that deserve concerted attention in the future. Much has been accomplished; much remains to be done.

References

- Alba, J.W., and A. Chattopadhyay (1985). Effects of context and part-category cues on recall of competing brands. *J. Marketing Res.* 22 (August), 340-349.
- Alba, J.W., and A. Chattopadhyay (1986). Saliency effects in brand recall. *J. Marketing Res.* 23 (November), 363-369.
- Alba, J.W., and J.W. Hutchinson (1987). Dimensions of consumer expertise. *J. Consumer Res.* 13, 411-454.
- Albaum, G., and D.I. Hawkins (1983). Geographic mobility and demographic and socioeconomic market segmentation. *J. Acad. Marketing Sci.* 11 (Spring), 110.

- Amemiya, T. (1981). Qualitative response models: A survey. *J. Econometric Literature* 19 (December), 1483-1583.
- Bagozzi, R.P. (1977). Structural equation models in experimental research. *J. Marketing Res.* 14 (May), 209-226.
- Bagozzi, R.P. (1980). Performance and satisfaction in an industrial salesforce: An examination of their antecedents and simultaneity. *J. Marketing* 44 (Spring), 65-77.
- Bagozzi, R.P. (1982). A field investigation of causal relations among cognitions, affect, intentions and behavior. *J. Marketing Res.* 19 (November), 562-584.
- Bass, F.M., and W.W. Talarzyk (1972). An attitude model for the study of brand preference. *J. Marketing Res.* 9 (February), 93-96.
- Batsell, R.R., and J.C. Polking (1985). A new class of market share models. *Marketing Sci.* 4(3), 177-198.
- Bearden, W.O., and M.J. Etzel (1982). Reference group influence on product and brand purchase decisions. *J. Consumer Res.* 9 (September), 183-194.
- Bearden, W.O., and J.E. Teel (1983). Selected determinants of consumer satisfaction and complaint reports. *J. Marketing Res.* 20 (February), 21-28.
- Bechtel, G.G. (1985). Generalizing the Rasch model for consumer rating scales. *Marketing Sci.* 4(1), 62-73.
- Bechtel, G.G. (1990). Share-ratio estimation of the nested multinomial logit model. *J. Marketing Res.* 27 (May), 232-237.
- Ben-Akiva, M., and S.R. Lerman (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*, Massachusetts Institute of Technology Press, Cambridge.
- Bettman, J.R. (1979). *An Information Processing Theory of Consumer Choice*, Addison-Wesley, Reading, MA.
- Biehal, G.J. (1983). Consumers' prior experiences and perceptions in auto repair choice. *J. Marketing* 47 (Summer), 82-91.
- Blattberg, R.C., and A.P. Jeuland (1981). A micro-modeling approach to investigate the advertising sales relationship. *Management Sci.* 27(9), 988-1004.
- Bodnar, J., P. Dilworth and S. Iacono (1988). Cross-sectional analysis of residential telephone subscription in Canada. *Inform. Econom. Policy* 3, 359-378.
- Bolton, R.N., and J.H. Drew (1991). A multistage model of customers' assessments of service quality and value. *J. Consumer Res.* 17 (March), 375-384.
- Bridges, S. (1991). A schema unification model of brand extensions, Babcock GSM Working Paper, Wake Forest University, Winston-Salem, NC.
- Bunch, D.S., and R.R. Batsell (1989). A Monte Carlo comparison of estimators for the multinomial logit model. *J. Marketing Res.* 26 (February), 56-68.
- Burnkrant, R.E., and T.J. Page, Jr. (1982). An examination of the convergent, discriminant, and predictive validity of Fishbein's behavioral intention model. *J. Marketing Res.* 19 (November), 550-561.
- Burt, R.S. (1987). Social contagion and innovation: Cohesion versus structural equivalence. *Amer. J. Sociology* 92 (May), 1287-1335.
- Carroll, D.J., and P.E. Green (1988). An INDSCAL-based approach to multiple correspondence analysis. *J. Marketing Res.* 25 (May), 193-203.
- Chapman, R.G., and R. Staelin (1982). Exploiting rank ordered choice set data within the stochastic utility model. *J. Marketing Res.* 19 (August), 288-301.
- Chatterjee, R., and J. Eliashberg (1990). The innovation diffusion process in a heterogeneous population: A micromodeling approach. *Management Sci.* 36(9), 1057-1074.
- Choffray, J.-M., and G.L. Lilien (1980). *Market Planning for New Industrial Products*, Wiley, New York.
- Churchill, G.A. Jr., and C. Suprenant (1982). An investigation into the determinants of customer satisfaction. *J. Marketing Res.* 19 (November), 491-504.
- Coleman, J.S., E. Katz and H. Menzel (1957). The diffusion of an innovation among physicians. *Sociometry* 20 (December), 253-270.
- Cooper, L.G. (1983). A review of multidimensional scaling in marketing research. *Appl. Psychol. Measurement* 7 (Fall), 427-450.
- Cooper, L.G., and M. Nakanishi (1983). Standardizing variables in multiplicative choice models. *J. Consumer Res.* 10 (June), 96-108.

- Corstjens, M.L., and D.A. Gautschi (1983). Formal choice models in marketing. *Marketing Sci.* 2(1), 19–56.
- Currim, I.S. (1982). Predictive testing of consumer choice models not subject to independence of irrelevant alternatives. *J. Marketing Res.* 19 (May), 208–222.
- Currim, I.S., and R.K. Sarin (1984). A comparative evaluation of multiattribute consumer preference models. *Management Sci.* 30(5), 543–561.
- Currim, I.S., and R.K. Sarin (1989). Prospect versus utility. *Management Science* 35(1), 22–41.
- Czepiel, J.A. (1974). Word of mouth processes in the diffusion of a major technological innovation. *J. Marketing Res.* 11 (May), 172–180.
- Daganzo, C. (1979). *Multinomial Probit*, Academic Press, New York.
- Dalal, S.R., and R.W. Klein (1988). A flexible class of discrete choice models. *Marketing Sci.* 7(3), 232–251.
- Day, G.S., and T. Deutscher (1982). Attitudinal predictions of choices of major appliance brands. *J. Marketing Res.* 19 (May), 192–198.
- DeGroot, M.H. (1970). *Optimal Statistical Decisions*, McGraw-Hill, New York.
- DeSarbo, W.S., and D.L. Hoffman (1987). Constructing MDS joint spaces from binary choice data: A multidimensional unfolding threshold model for marketing research. *J. Marketing Res.* 24 (February), 40–54.
- DeSarbo, W.S., A.K. Manrai and L.A. Manrai (1993). Non-spatial tree models for the assessment of competitive market structure: An integrated review of the marketing and psychometric literature, in: J. Eliashberg and G.L. Lilien (eds.), *Handbooks in Operations Research and Management Science, Vol. 5: Marketing*, North-Holland, Amsterdam, ch. 5.
- Dick, A., D. Chakravarti and G. Biehal (1990). Memory-based inferences during consumer choice. *J. Consumer Res.* 17 (June), 82–93.
- Domencich, T.A., and D. McFadden (1975). *Urban Travel Demand: A Behavioral Analysis*, North-Holland, Amsterdam.
- Dubin, J.A. (1986). A nested logit model of space and water heat system choice. *Marketing Sci.* 5(2), 112–124.
- Ehrenberg, A.S.C., G.J. Goodhardt and T.P. Barwise (1990). Double jeopardy revisited. *J. Marketing* 54(3), 82–91.
- Eliashberg, J. (1980). Consumer preference judgements: An exposition with empirical applications. *Management Sci.* 26(1), 60–77.
- Eliashberg, J., and J.R. Hauser (1985). A measurement error approach for modelling consumer risk preferences. *Management Sci.* 31(1), 1–25.
- Elrod, T. (1988). Choice map: Inferring a product-market map from panel data. *Marketing Sci.* 7(1), 21–40.
- Farquhar, P.H. (1984). Utility assessment methods. *Management Sci.* 30(11), 1283–1300.
- Fishbein, M. (1963). An investigation of relationship between beliefs about an object and the attitude toward the object. *Human Relations* 16, 233–240.
- Fishbein, M. (1967). Attitude and prediction of behavior, in M. Fishbein (ed.), *Readings in Attitude Theory and Measurement*, Wiley, New York, pp. 477–492.
- Fishbein, M., and I. Ajzen (1975). *Beliefs, Attitude, Intention and Behavior*, Addison-Wesley, Reading, MA.
- Fishburn, P.C. (1984). Multiattribute nonlinear utility theory. *Management Sci.* 30(11), 1301–1310.
- Fotheringham, A.S. (1988). Consumer store choice and choice set definition. *Marketing Sci.* 7(3), 299–310.
- Gensch, D.H. (1984). Targeting the switchable industrial customer. *Marketing Sci.* 3(1), 41–54.
- Gensch, D.H. (1985). Empirically testing a disaggregate choice model for segments. *J. Marketing Res.* 22 (November), 462–467.
- Gensch, D.H. (1987a). Empirical evidence supporting the use of multiple choice models in analyzing a population. *J. Marketing Res.* 24 (May), 197–207.
- Gensch, D.H. (1987b). A two-stage disaggregate attribute choice model. *Marketing Sci.* 6(3), 223–239.
- Gensch, D.H., N. Aversa and S.P. Moore (1990). A choice modeling market information system that enabled ABB Electric to expand its market share. *Interfaces* 20(1), 6–25.
- Gensch, D.H., and W.W. Recker (1979). The multinomial, multiattribute logit choice model. *J. Marketing Res.* 16 (February), 124–132.

- Gensch, D.H., and J.A. Svestka (1984). A maximum likelihood hierarchical disaggregate model for predicting choice of individuals. *J. Math. Psychol.* 28, 160-178.
- Gessner, G., W.A. Kamakura, N.K. Malhotra and M.E. Zmijewski (1988). Estimating models with binary dependent variables: Some theoretical and empirical observations. *J. Business Res.* 16(1), 49-65.
- Green, P.E., S.M. Goldberg and M. Montemayor (1981). A hybrid utility estimation model for conjoint analysis. *J. of Marketing* 45 (Winter), 33-41.
- Green, P.E., and V.R. Rao (1972). *Applied Multidimensional Scaling*, Holt, Rinehart and Winston, New York.
- Green, P.E., and V. Srinivasan (1978). Conjoint analysis in consumer research: Issues and outlook. *J. Consumer Res.* 5 (September), 103-123.
- Green, P.E., and Y. Wind (1973). *Multi-Attribute Decisions in Marketing*, Dryden Press, Hinsdale, IL.
- Guadagni, P.M., and J.D.C. Little (1983). A logit model of brand choice calibrated on scanner data. *Marketing Sci.* 2(3), 203-238.
- Guadagni, P.M., and J.D.C. Little (1987). When and what to buy; A nested logit model of coffee purchase, Working Paper #1919-87, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA (August).
- Gupta, S. (1988). Impact of sales promotion on when, what, and how much to buy. *J. Marketing Res.* 25 (November), 342-355.
- Hagerty, M.R., and D.A. Aaker (1984). A normative model of consumer information processing. *Marketing Sci.* 3(3), 227-246.
- Harshman, R.A., P.E. Green, Y. Wind and M.E. Lundy (1982). A model for the analysis of asymmetric data in marketing research. *Marketing Sci.* 1(2), 205-242.
- Hauser, J.R. (1986). Agendas and consumer choice. *J. Marketing Res.* 23 (August), 199-212.
- Hauser, J.R., and F.S. Koppelman (1979). The relative accuracy and usefulness of alternative perceptual mapping techniques. *J. Marketing Res.* 14(3), 495-507.
- Hauser, J.R., J.H. Roberts and G.L. Urban (1983). Forecasting sales of a new durable, in F.S. Zufryden (ed.), *Advances and Practices of Marketing Science Management*, The Institute of Management Science, Providence, RI, pp. 115-128.
- Hauser, J.R., and S.M. Shugan (1980). Intensity measures of consumer preference. *Oper. Res.* 28(2), 278-320.
- Hauser, J.R., and P.D. Simmie (1981). Profit maximizing product positioning. *Management Sci.* 27(1), 33-56.
- Hauser, J.R., and G.L. Urban (1979). Assessment of attribute importance and consumer utility functions: Von Neumann-Morgenstern theory applied to consumer behavior. *J. Consumer Res.* 5 (March), 251-262.
- Hauser, J.R., and G.L. Urban (1986). The value priority hypotheses for consumer budget plans. *J. Consumer Res.* 12 (March), 449-462.
- Hauser, J.R., and B. Wernerfelt (1990). An evaluation cost model of consideration sets. *J. Consumer Res.* 16 (March), 393-408.
- Hauser, J.R., and K.J. Wiesniewski (1982). Dynamic analysis of consumer response to marketing strategy. *Management Sci.* 28(5), 455-486.
- Hoffman, D.L., and G.R. Franke (1986). Correspondence analysis: Graphical representation of categorical data in marketing research. *J. Marketing Res.* 23 (August), 213-227.
- Horsky, D., and M.R. Rao (1984). Estimation of attribute weights from preference comparisons. *Management Sci.* 30(7), 801-822.
- Howard, J.A., and J.N. Sheth (1969). *The Theory of Buyer Behavior*, Wiley, New York.
- Huber, J. and M.B. Holbrook (1979). Using attribute ratings for product positioning: Some distinctions among compositional approaches. *J. Marketing Res.* 16 (November), 507-516.
- Huber, J., and J. McCann (1982). The impact of inferential beliefs on product evaluations. *J. Marketing Res.* 19 (August), 324-333.
- Jain, D.C., and N.J. Vilcassim (1991). Investigating household purchase timing decisions: A conditional hazard function approach. *Marketing Sci.* 10(1), 1-23.
- Johnson, E.J., and R.J. Meyer (1984). Compensatory choice models of noncompensatory processes: The effect of varying context. *J. Consumer Res.* (June), 528-541.

- Johnson, E.J., R.J. Meyer and S. Ghose (1989). When choice models fail: Compensatory models in negatively correlated environments. *J. Marketing Res.* 26 (August), 255-270.
- Johnson, M.D. (1984). Consumer choice strategies for comparing noncomparable alternatives. *J. Consumer Res.* 11 (December), 741-753.
- Johnson, M.D. (1986). Modeling choice strategies for noncomparable alternatives. *Marketing Sci.* 5(1), 37-54.
- Johnson, M.D. (1988). Comparability and hierarchical processing in multialternative choice. *J. Consumer Res.* 15 (December), 303-314.
- Jones, J.M., and F.S. Zufryden (1981). Relating deal purchases and consumer characteristics to repeat purchase probability. *J. Marketing Res. Soc.* 23, 84-99.
- Jöreskog, K., and D. Sörbom (1979). *Advances in Factor Analysis and Structural Equation Models*, Abt Books, Cambridge, MA.
- Kahn, B.E., M.U. Kalwani and D.G. Morrison (1986). Measuring variety-seeking and reinforcement behaviors using panel data. *J. Marketing Res.* 23 (May), 89-100.
- Kahneman, D., and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica* 47 (March), 263-291.
- Kamakura, W.A., and R.K. Srivastava (1984). Predicting choice shares under conditions of brand interdependence. *J. Marketing Res.* 21 (November), 420-434.
- Kamakura, W.A., and R.K. Srivastava (1986). An ideal-point probabilistic choice model for heterogeneous preferences. *Marketing Sci.* 5(3), 199-218.
- Katahira, H. (1990). Perceptual mapping using ordered logit analysis. *Marketing Sci.* 9(1), 1-17.
- Keeney, R.L., and H. Raiffa (1976). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, Wiley, New York.
- Keller, K.L., and R. Staelin (1989). Assessing biases in measuring decision effectiveness and information overload. *J. Consumer Res.* 15 (March), 504-508.
- Lancaster, K. (1966). A new approach to consumer theory. *J. Political Economy* 74(2), 132-157.
- Lattin, J.M. (1987). A model of balanced choice behavior. *Marketing Sci.* 6(1), 48-65.
- Lattin, J.M., and R.E. Bucklin (1989). Reference effects of price and promotion on brand choice behavior. *J. Marketing Res.* 26 (August), 299-310.
- Lattin, J.M., and L. McAlister (1985). Using a variety-seeking model to identify substitute and complementary relationships among competing products. *J. Marketing Res.* 22 (August), 330-339.
- Lattin, J.M., and J.H. Roberts (1988). Modeling the role of risk-adjusted utility in the diffusion of innovation, Research paper 1019, Graduate School of Business, Stanford University, Stanford, CA (October).
- Leckenby, J.D., and S. Kishi (1984). The Dirichlet multinomial distribution as a magazine exposure model. *J. Marketing Res.* 21 (February), 100-106.
- Lehmann, D.R. (1971). Television show preference: Application of a choice model. *J. Marketing Res.* 8 (February), 47-55.
- Lilien, G.L., P. Kotler and K.S. Moorthy (1992). *Marketing Models*, Prentice-Hall, Englewood Cliffs, NJ.
- Lindblom, C.E. (1959). The science of muddling through. *Public Admin. Rev.* 19, 79-88.
- Lisco, T. (1967). The value of commuter's travel time: A study of urban transportation, Ph.D. dissertation, Department of Economics, University of Chicago.
- Long, J.S. (1983). *Confirmatory Factor Analysis*. Quantitative Applications in the Social Sciences, Series No. 33, Sage Publications, Beverly Hills, CA.
- Louviere, J.J., and D.A. Hensher (1983). Using discrete choice models with experimental design data to forecast consumer demand for a unique cultural event. *J. Consumer Res.* 10 (December), 348-361.
- Louviere, J.J., and G. Woodworth (1983). Design and analysis of simulated consumer choice or allocated experiments: An approach based on aggregate data. *J. Marketing Res.* 20 (November), 350-367.
- Luce, R.D. (1959). *Individual Choice Behavior*, Wiley, New York.
- Lynch, J.G., Jr., and T. Srull (1982). Memory and attentional factors in consumer choice: Concepts and research methods. *J. Consumer Res.* 9 (June), 18-37.
- MacKay, D.B., and J.L. Zinnes (1986). A probabilistic model for the multidimensional scaling of proximity and preference data. *Marketing Sci.* 5(4), 325-344.
- Mahajan, V., E. Muller and S. Sharma (1984). An empirical comparison of awareness forecasting models of new product acceptance. *Marketing Sci.* 3 (Summer), 179-197.

- Malhotra, N.K. (1982). Information load and consumer decision making. *J. Consumer Res.* 8 (March), 419-430.
- Malhotra, N.K. (1986). An approach to the measurement of consumer preferences using limited information. *J. Marketing Res.* 23 (February), 33-40.
- Manrai, A.K., and P. Sinha (1989). Elimination-by-cutoffs. *Marketing Sci.* 8(2), 133-152.
- Mansfield, E. (1961). Technical change and the rate of imitation. *Econometrica* 29(4), 741-765.
- Massy, W.F., D.B. Montgomery and D.G. Morrison (1970). *Stochastic Models of Buying Behavior*, Massachusetts Institute of Technology Press, Cambridge.
- McAlister, L. (1982). A dynamic attribute satiation model of variety-seeking behavior. *J. Consumer Res.* 9 (September), 141-150.
- McAlister, L., and E. Pessemier (1982). Variety-seeking behavior: An interdisciplinary review. *J. Consumer Res.* 9 (December), 311-322.
- McFadden, D. (1976). Quantal choice analysis: A survey. *Ann. Econom. Social Measurement* 5 (May), 363-369.
- McFadden, D. (1978). Modeling to choice of residential location, in A. Karlquist et al. (eds.), *Spatial Interaction Theory and Residential Location*, North-Holland, Amsterdam.
- McFadden, D. (1980). Econometric models for probabilistic choice among products. *J. Business*, 53(3), pt. 2 (July), 513-530.
- McFadden, D. (1986). The choice theory approach to market research. *Marketing Sci.* 5(4), 275-297.
- McFadden, D. (1991). Advances in computation, statistical methods and testing of discrete choice models. *Marketing Lett.* 2(3), 215-230.
- Meyer, R.J., and T.C. Eagle (1982). Context-induced parameter instability in a disaggregate-stochastic model of store choice. *J. Marketing Res.* 19 (February), 62-71.
- Meyer, R.J., and A. Sathi (1985). A multiattribute model of consumer choice during product learning. *Marketing Sci.* 4(1), 41-61.
- Midgley, D.F., P.D. Morrison and J.H. Roberts (1992). The effect of network structure in industrial diffusion processes. *Res. Policy* 21(6) 533-552.
- Moore, W.L., and R.S. Winer (1987). A panel-data based method for merging joint space and market response function estimation. *Marketing Sci.* 6(1), 25-42.
- Mukherjee, B.N. (1973). Analysis of covariance structures and exploratory factor analysis. *Brit. J. of Math. Statist. Psychol.* 26 (November), 125-154.
- Nedungadi, P. (1990). Recall and consumer consideration sets: Influencing choice without changing brand evaluations. *J. Consumer Res.* 17(3), 263-276.
- Nooteboom, B. (1989). Diffusion, uncertainty and firm size. *Intern. J. Res. Marketing* 6, 109-128.
- Oliver, R.L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *J. Marketing Res.* 17 (November), 460-469.
- Oliver, R.L., and W.S. DeSarbo (1988). Response determinants in satisfaction judgments. *J. Consumer Res.* 14 (March), 495-507.
- Oren, S.S. and R.G. Schwartz (1988). Diffusion of new products in risk-sensitive markets. *J. Forecasting* 7 (October/December), 273-287.
- Papatla, P., and L. Krishnamurthi (1992). A probit model of choice dynamics. *Marketing Sci.* 11(2) 189-206.
- Parasuraman, A., V.A. Zeithaml and L.L. Berry (1985). A conceptual model of service quality and its implications for future research. *J. Marketing* 49(4), 41-50.
- Payne, J.W., J.R. Bettman and E.J. Johnson (1988). Adaptive strategy selection in decision making. *J. Experimental Psychol.* 14, 534-552.
- Punj, G.N., and R. Staelin (1978). The choice for graduate business schools. *J. Marketing Res.* 15 (November), 588-598.
- Ratchford, B.T. (1982). Cost-benefit models for explaining consumer choice and information seeking behavior. *Management Sci.* 28(2), 197-212.
- Richins, M.L. (1983). Negative word-of-mouth by dissatisfied consumers: A pilot study. *J. Marketing* 47 (Winter), 68-78.
- Roberts, J.H., and J.M. Lattin (1991). Development and testing of a model of consideration set composition. *J. Marketing Res.* 28 (November), 429-440.

- Roberts, J.H., and G.L. Urban (1988). Modeling multiattribute utility, risk, and belief dynamics for new consumer durable brand choice. *Management Sci.* 34(2), 167-185.
- Robinson, M. (1985). A logit model of the all electric residence decision. *Rev. Regional Econom. Business* 11(2), 17-20.
- Rogers, E.M. (1983). *Diffusion of Innovations*, 3rd edition, The Free Press, New York.
- Rossiter, J.R., and L. Percy (1987). *Advertising and Promotion Management*, McGraw-Hill, New York.
- Sheth, J.N. (1974). *Models of Buyer Behavior: Conceptual, Quantitative, and Empirical*, Harper and Row, New York.
- Shocker, A.D., and V. Srinivasan (1979). Multi-attribute approaches for product concept evaluation and generation: A critical review. *J. Marketing Res.* 16 (May), 159-180.
- Shugan, S.M. (1980). The cost of thinking. *J. Consumer Res.* 7 (September), 99-111.
- Simonson, I. (1990). The effect of purchase quantity and timing on variety-seeking behavior. *J. Marketing Res.* 27 (May), 150-162.
- Singh, J. (1988). Consumer complaint intentions and behavior: Definitional and taxonomical issues. *J. Marketing* 52 (January), 93-107.
- Stigler, G. (1961). The economics of information. *J. Political Economy* 69(3), 213-225.
- Sujan, M. (1985). Consumer knowledge: Effects on evaluation strategies mediating consumer judgments. *J. Consumer Res.* 12, 31-46.
- Swan, J.E., and L.J. Combs (1976). Product performance and consumer satisfaction: A new concept. *J. Marketing* 40 (April), 25-33.
- Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Sci.* 4(3), 199-214.
- Tse, D.K., and P.C. Wilton (1988). Models of consumer satisfaction formation: An extension. *J. Marketing Res.* 25 (May), 204-212.
- Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychol. Rev.* 79, 281-299.
- Tversky, A., and D. Kahneman (1974). Judgment under uncertainty: Heuristics and biases. *Science* 185, 1124-1131.
- Urban, G.L., and J.R. Hauser (1980). *Design and Marketing of New Products*, Prentice-Hall, Englewood Cliffs, NJ.
- Urban, G.L., J.R. Hauser and J.H. Roberts (1990). Prelaunch forecasting of new automobiles. *Management Sci.* 36(4), 401-421.
- Urban, G.L., J.S. Hulland and B.D. Weinberg (1990). Modeling categorization, elimination, and consideration for new product forecasting of consumer durables, Working Paper No. 3206-90, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA.
- Urbany, J.E., T.J. Madden and P.R. Dickson (1989). All's not fair in pricing: An initial look at the dual entitlement principle. *Marketing Lett.* 1(1), 17-25.
- Wagner, U., and A. Taudes (1986). A multivariate Polya model of brand choice and purchase incidence. *Marketing Sci.* 5(3), 219-244.
- Westbrook, R.A. (1987). Product/consumption-based affective responses and postpurchase processes. *J. Marketing Res.* 24 (August), 258-270.
- Wilson, D.T., H.L. Matthews and J.W. Harvey (1975). An empirical test of the Fishbein behavioral intention model. *J. Consumer Res.* 1, 39-48.
- Winer, R.S. (1986). A reference price model of brand choice for frequently purchased products. *J. Consumer Res.* 13 (September), 250-256.
- Woodruff, R.B., E.R. Cadotte and R.L. Jenkins (1983). Modeling consumer satisfaction processes using experience-based norms. *J. Marketing Res.* 20 (August), 296-304.
- Wright, P.L., and F. Barbour (1977). Phased decision strategies: Sequels to an initial screening, in M. Starr and M. Zeleny (eds.), *Multiple Criteria Decision Making, North-Holland TIMS Studies in the Management Sciences*, Vol. 6, North-Holland, Amsterdam, pp. 91-109.
- Zaltman, G., and M. Wallendorf (1979). *Consumer Behavior: Basic Findings and Management Implications*, Wiley, New York.
- Zeithaml, V.A., A. Parasuraman and L.L. Berry (1990). *Delivering Quality Service: Balancing Customer Perceptions*, The Free Press, New York.