"A NEW BRAND CHOICE MODEL INCORPORATING A CHOICE SET FORMATION PROCESS"

by

W. VANHONACKER*

94/03/EPS/MKT

* Professor of Marketing at INSEAD, Boulevard de Constance, Fontainebleau 77305 Cedex, France.

Printed at INSEAD, Fontainebleau, France

A New Brand Choice Model Incorporating A Choice Set Formation Process

Wilfried R. Vanhonacker *

Revised, December 1993

* Professor of Marketing, INSEAD, Fontainebleau, France and the Hong Kong University of Science & Technology, Hong Kong. This research is part of a single-source model development project for A.C. Nielsen and Euromedia (France). The paper has benefited from the comments of Niraj Dawar, Lydia Price, and two anonymous reviewers.

A New Brand Choice Model Incorporating
A Choice Set Formation Process

ABSTRACT

This paper develops an individual-level choice model which embeds a rational choice set generation process. The choice set generation process is motivated by a goal-derived categorization paradigm where consumers make uncertain judgments about the ability of brands to meet the consumption or usage goal motivating the choice decision. Recognizing this uncertainty, brand selection is modeled as a decision under uncertainty within a random utility framework. The modeling of the process builds on previously suggested formalizations based on search-theory and decision costs. Derivations are provided for a flexible and parsimonious representation of the choice probabilities which are characterized neither by IIA nor the regularity property. Some initial empirical results using scanner data are reported.

Key Words: **Discrete Choice Models; Choice Set Generation; Goal-Derived Categorization; Typicality**

INTRODUCTION

Parsimonious models of the individual consumer's choice process have an extensive tradition in marketing. Describing choice as a rational process has helped in our understanding and explanation of observed choice outcomes. The utility maximizing choice rule has played a central role in that regard. The rational process of utility maximization leads the individual consumer to choose the alternative with the highest utility. Hence, choice results from a relative comparison conditional on a choice set which consists of all available alternatives actually considered. Some authors have referred to this set as the consideration set after Wright and Barbour (1977). Choice set generation then refers to the process of defining the "considered subset' (Simon 1955).

Until recently, little attention has been devoted in the marketing brand choice literature to the process by which such choice sets are formed. Indeed, most operationalizations define choice as the selection of a particular brand from all those available in the market place (see, e.g. Gensch and Recker 1979, Guadagni and Little 1983, Gupta 1988, Lattin and Bucklin 1989, Chintagunta, Jain and Vilcassim 1991, Fader, Lattin and Little 1993, Fader and Lattin 1993). This tradition of equating choice set with the set of all available alternatives is rather surprising given conflicting behavioral evidence which indicates that consumers often select from a limited number of items (see, e.g., Wright and Barbour 1977, Bettman 1979, 1987, Payne 1982). This disparity between choice models and empirical evidence is troubling in light of extensive empirical and theoretical work in the transportation literature which argues that the implications of choice set misspecification can be quite severe (see, e.g., Stopher 1980, Williams and Ortuzar 1982, Swait 1984, Swait and Ben-Akiva 1986).

Recognizing these dangers, recent work in marketing has followed the transportation literature (see, e.g., Ben-Akiva 1977, Swait 1984, Pitschke 1980, Swait and Ben-Akiva 1987, Boccara 1989, Ben-Akiva and Boccara 1990) in focusing on the issue of choice or consideration set generation (Roberts 1989, Nedungadi 1987, 1990, Shocker et. al. 1991), and initial attempts to integrate the choice set generation process into discrete consumer choice models have appeared (Roberts and Lattin 1991, Hauser and Wernerfelt 1989, 1990). Hauser and Wernerfelt (1989, 1990), for example, formalize the process of choice set formation in a rational fashion using search theory (analogous to Meyer 1982 and Richardson 1982). Building on the information search paradigm, Roberts and Lattin (1991) provide an operationalization, similar to that of Ben-Akiva and Lerman (1985), of the incremental benefit derived from including another alternative in the choice set. This

representation of choice set generation, which is calibrated separately across consumers, is integrated into a multinomial logit model of choice.

While improving on earlier models which completely lack a choice set specification, these latter modeling efforts still fail to recognize an important aspect of the choice set generation process. As discussed in Shocker et. al. (1991), the usage or consumption occasion motivating the choice is likely to influence the nature and composition of the choice set. Indeed, Ratneshwar and Shocker (1991) argue that usage contexts act as environmental constraints that help the individual consumer define ends and goals, and, hence, limit the nature of the brands that can help achieve consumption objectives. The notion of consumption objectives, or goals, implies that the consumer will judge the extent to which brands will enable goal achievement. This judgment is central to the choice set generation process.

One approach which recognizes the inherent characteristics of the choice set generation process is to conceptualize it as a goal-derived categorization process (Barsalou 1983, 1985). The goal-derived categorization paradigm, invoking the construct of typicality, has been well accepted in consumer behavior (see, e.g., Troye 1984, Sujan 1985, Nedungadi and Hutchinson 1985, Ward and Loken 1986, Loken and Ward 1987, Nedungadi 1990, Nedungadi and Kanetkar 1992). Typicality refers to the phenomenon that not all members of a category are equally representative of that category. In goal-derived categorization, typicality of an object depends on its ability to meet a goal, with ability varying across category members. Choice set generation can be viewed as a categorization task with the usage or consumption occasion motivating the choice as the goal. Here, typicality would refer to the ability of the choice set members to help achieve the consumption or usage goal. As there is evidence that consumers organize categories around typical or exemplar brands (Sujan 1985), choice set formation then arises from typicality judgments relative to a "prototypical" or "ideal" brand for the consumption goal.

In this paper, the choice set generation process is conceptualized as a goal-derived categorization process. A rational operationalization of the process is provided which explicitly recognizes possible uncertainty in typicality judgments and, hence, choice set membership. This operationalization is integrated in a choice model which combines the random utility evaluations of the brands typical in traditional discrete choice models with their uncertain typicality evaluations given a particular usage or consumption occasion. As a rational individual will attempt to resolve the uncertainty (and, hence, the risk associated with each brand) in the brand selection task, the choice rule is one of maximizing expected utility. The derived choice model is parsimonious and is not characterized by either the Independence of Irrelevant Alternatives (IIA) property (Currim 1982) or the regularity property (Huber, Payne, and Puto, 1982).

The contribution of the choice model developed here lies in its comprehensive integration of a behaviorally-motivated choice set generation process in a random utility choice framework. It illustrates the importance of considering the usage context or consumption occasion motivating the brand choice decision and its impact on the rational choice rule. Conceptually, the model combines theories of goal-derived categorization in cognitive psychology with theories of random utility and decisions under uncertainty. Operationally, the model combines inherent uncertainty in choice set generation with random utility evaluation inherent in the choice task within an overall rational framework.

The paper is structured as follows. First, the theoretical background of the choice set generation process is discussed. Second, the basic model structure is derived in a random utility framework. Third, the choice set generation process is operationalized. Fourth, the choice model and its characteristics are discussed. Fifth, some empirical examples using scanner data are reported. Conclusions and directions for future research follow.

BACKGROUND

In discrete choice modeling the individual consumer is generally considered to select a brand out of a choice set which is a subset of the universal set containing all brands. The rationale is that the consumer has neither the ability nor the option to consider all brands in the universal set. Furthermore, not all brands might be judged appropriate for the consumption or usage occasion the consumer has in mind. The question arises as to how the individual consumer will decide on the composition of the choice set. No doubt, external constraints such as budget and availability will play a role. Of interest here is to explore the internal evaluation process the consumer will follow alongside the external constraints to compose the choice set.

Consider an individual consumer who feels somewhat hungry at mid-afternoon and is looking for a snack. Upon doctor's advice, he has recently started to change his eating habits to control his body weight. Consequently, he prefers a light, healthy snack. Because of calorie considerations and natural ingredients, items such as an apple, a yogurt, and a granola bar come to mind. Feeling guilty about wanting to eat something, the consumer considers plain yogurt as the least harmful although he has never developed a liking for yogurt. As he is fond of apricots, he decides to limit his choice to apricot-flavored yogurts. As yogurts are stacked by brand and not by flavor on the shelf, he looks

at what is available and notices Danone, his wife's favorite brand. Knowing little about yogurt but remembering his wife's health consciousness, he compares the available apricot flavors to Danone apricot. Yoplait seems to have similar ingredients, but as it comes in a slightly smaller pot he selects apricot-flavored Yoplait.

Although the example is specific, it is neither unrealistic nor uncommon as to the process of identifying a choice set and the selection of an alternative from that set. As the example illustrates, the choice set generation process is one of sequentially evaluating items for membership and it involves a complex interaction between the individual consumer, the consumption or usage goal, and the alternatives. In the above example, the process is influenced by the objective of selecting a light, healthy snack, by the consumer's preference for apricot, by the consideration of plain yogurt as being an ideal snack given the consumption objective, and by his lack of familiarity and experience with any particular brand.

The choice set generation process is a categorization process based on the objective of realizing a goal. In the example, the goal is the consumption of a light, healthy snack. Sequential evaluation and categorization of items is driven by that goal. In psychology, one refers to such a process as goal-derived categorization (Barsalou 1983, 1985). Cognitive psychologists have studied the process primarily at the product level (i.e., apples versus cookies versus yogurt) (see, e.g., Rosch 1973, 1975, Rosch and Mervis 1975, Smith and Melara 1990, Fried and Holyoak 1984). Consumer behaviorists have focused on the categorization process at the brand level where items to be categorized are more homogeneous in terms of descriptive attributes (i.e., Danone yogurt versus Yoplait yogurt) (see, e.g., Troye 1984, Sujan 1985, Nedungadi and Hutchinson 1985, Ward and Loken 1986, Loken and Ward 1987, Nedungadi 1990, Anderson 1991, Nedungadi and Kanetkar 1992). The categorization paradigm at the brand level has proven useful in research on brand extensions (see, e.g., Herr et. al. 1990, Boush and Loken 1991, Dawar and Anderson 1993) and pioneering advantages (see, e.g., Carpenter and Nakamoto 1989, Carpenter et. al. 1993).

Central to the categorization paradigm is the concept of typicality which refers to the phenomenon that not all members of a category are equally representative of that category. In goal-derived categorization, typicality refers to the extent to which an alternative is judged to help achieve the goal. In our example, plain yogurt is more typical of a light, healthy snack then a granola bar. Indeed, the consumer judges plain jogurt as the "prototypical" alternative given the consumption goal. Other options such as apricot-flavored yogurt, an apple, etc. are considered increasingly atypical. At the brand level, the consumer in our

example considers Danone the prototypical brand. The categorization is carried out via comparisons of items to a prototypical item and is captured in degrees of typicality on a continium going from prototypical to atypical. In essence, the prototypical item represents an ideal given the consumption goal.

Extensive research has been done on the determinants of typicality. Cognitive psychologists, who have focused primarily on the categorization of distinct items, have identified the influence of feature similarity (or family resemblance) and goal achievement (ideals, attribute structure) (see, e.g., Rosch and Mervis 1975, Barsalou 1983, 1985). Consumer behaviorists, who have focused on categorization at the brand level where items are more homogeneous in terms of descriptive attributes, have identified attitude towards items as an important determinant (see, e.g., Loken and Ward 1990). This would suggest that in the categorization process, as items become more "similar", preferences will increasingly influence ideosyncractic typicality judgments. Where in the above example an apple and a yogurt are considered typical healthy foods, perhaps because of lack of calories or artificial ingredients, the dislike for plain yogurt becomes important despite it being a prototypical healthy snack. Note, however, that not all consumer behavior theories support a positive link between preference and typicality. Variety seeking and desire for uniqueness and innovativeness (Snyder and Fromkin 1979) would imply strong preference for what is atypical. Given that greater familiarity leads to greater liking (Zajonic 1968), familiarity and experience can be expected to be a strong mediating factor in the relationship between preference and typicality.

The importance of preference as a determinant signals that typicality judgments at the brand level are strongly influenced by perceived performance judgments on valued dimensions. Those could be quite abstract and intangible. Consistent with research on cognitive representation (see, e.g., Johnson and Fornell 1987), those judgments are probably spatial (i.e., multidimensional). Spatial evaluations on continuous dimensions imply relative typicality represented in relative distances to the prototypical item. More importantly, they suggest that choice set members derived from the goal-derived categorization process are not equally substitutable. In the above example, Danone is the prototypical brand. Other brands are judged at different "distances" from Danone and their pairwise substitutability depends on their relative distance to the prototypical brand. This is in sharp contrast to the treatment of choice alternatives belonging to the choice set as equally substitutable in traditional discrete choice models (Ben-Akiva and Lerman 1985).

Given the mediating effect of familiarity, typicality judgments are characterized by uncertainty. In our example, the consumer has limited familiarity with alternatives as well

as the consumption goal. Only having recently started worrying about his health, he is uncertain about various snacks and their potential impact on his health. Apart from Bayesian learning models in cognitive psychology (see, e.g., Fried and Holyoak 1984), little attention has been devoted to this uncertainty in the categorization literature. In studying choice set generation processes and particularly their integration into the choice task, consideration of this uncertainty is very important as the individual consumer will address and attempt to resolve it within the choice task.

In general, the uncertainty in typicality judgments implies uncertainty about what is prototypical and uncertainty about choice set membership. Uncertainty about which item is prototypical might lead the consumer to consider several potentially prototypical items jointly instead of considering a single item. The prototypicality contained in the set would then function as a cognitive reference point against which other items would be judged in terms of typicality.

Uncertainty about choice set membership implies that each item has some likelihood of inclusion. Furthermore, any discrete boundary of the choice set is arbitrary with some items included despite being relatively atypical and others excluded despite being relatively typical.

Both aspects of uncertainty imply that the consumer is faced with the risk of selecting an item out of the choice set that might not help in achieving the consumption goal. This risk increases with items that are more atypical (and, hence, are closer to the boundaries of the choice set). This risk will affect the choice rule the consumer will use to select an item from the choice set. As has been established in the literature on risky choice (decision under uncertainty), the uncertainty attached to each alternative will be explicitly incorporated in the choice task. A flexible modeling approach is needed which recognizes these inherent characteristics of choice set generation and their impact on the choice task itself.

BASIC CHOICE MODEL STRUCTURE

Following Nedungadi (1990) who argues that choice set generation and brand selection play different roles, both are modeled separately here. The choice set generation process is modeled as a goal-derived categorization process as conceptualized above. The evaluation of items belonging to the choice set is modeled in a typical random utility framework. Both choice set generation and brand selection are viewed as rational processes which possibly

share identical determinants. Both processes imply uncertainty, but those uncertainties are considered to be independent with the uncertainty inherent in the categorization process being resolved within the choice rule characterizing the brand selection task. This is discussed in some detail now.

In the tradition of random utility models, the utility that an individual consumer i attaches to brand j can be expressed as

$$U_{ij} = \sum_{l} a_{ijl}\, f(X_{ijl}) + \mu_{ij}$$

or, more compactly,

$$= V_{ij} + \mu_{ij} \tag{1}$$

where $V_{ij} = \sum_{l} a_{ijl}\, f(X_{ijl})$ denotes the deterministic utility component, and μ_{ij} denotes the random utility component. The deterministic component V_{ij} is modeled as linear additive in functions of l explanatory variables, X_{ijl}. These functions are denoted by $f(X_{ijl})$ and the a_{ijl}'s represent the parameters of these functions in the linear additive specification. Model (1) is the fundamental random utility hypothesis underlying a large number of rational choice models (Corstjens and Gautschi 1983).

The utility-maximizing choice rule implies that brands are compared on a compensatory basis rather than on the basis of individual components (i.e. the X_{ijl}'s in the deterministic utility component V_{ij}) as is the case in other (conjunctive) choice rules (for a review on compensatory models, see Johnson and Meyer 1984). Given this choice rule, the probability that individual consumer i selects brand j out of a choice set M equals $P_{ij} = \text{Prob.}\left[U_{ij} > U_{ik}\,(k \in J, j \neq k)\right]$. With the assumption that the random utility components μ_{ij} are independent and identically distributed (i.i.d.) across brands $j \in M$, this probability can be expressed as (McFadden 1974),

$$P_{ij} = \int_{x=-\infty}^{+\infty} g(\mu_{ij} = x) \prod_{\substack{k \in J \\ k \neq j}} \int_{y=-\infty}^{V_{ij} - V_{ik} + x} g(\mu_{ij} = y).\, dy.\, dx$$

where g() represents the probability density function of the random component of utility in (1).

In traditional discrete choice modeling, the composition of choice set M is assumed to be either known or deterministically predictable (Ben-Akiva and Lerman 1985). There is no uncertainty about its composition. This is not the case when choice sets arise from goal-derived categorization. As discussed above, typicality judgments with their inherent uncertainty will give rise to a probability or likelihood of each item in the universal set (i.e., the set of all available brands, Shocker et. al. 1991) belonging to choice set M. As this uncertainty implies a risk of potentially selecting an item which might not help in achieving the consumption goal, the consumer is likely to integrate that uncertainty into the selection decision.

Choice among risky alternatives has been studied extensively in the decision theory literature (see, e.g., Schoemaker 1982, Machina 1987, Fishburn 1988). The choice task is generally framed as a choice among lotteries with specific monetary outcomes (either gains or losses). The lottery concept suggests that a certain gain or loss might not be obtained, and in essence is analogous to the uncertainty implied in choice set membership. An atypical item will have a low probability of belonging to the choice set implying a low likelihood of helping to achieve the consumption goal. In terms of risky choice terminology, brand selection then is a choice between lotteries with each brand essentially being a lottery with a payout equal to the utility of that brand (as captured in (1)) and the probability of that payout occurring equal to the probability of that brand belonging to the choice set. The only difference between this problem and the typical decision-under-uncertainty problem is the random payout. In a typical lottery, the monetary gains and losses are exact and known. In brand selection, the utilities of the brands are not known exactly. Moreover, brand selection is a choice among lotteries with uncertain outcomes.

Of importance here is the choice rule used in risky choice. The traditional paradigm has been one of maximizing expected utility (Schoemaker 1982). In other words, the most preferred lottery would be the one for which the expected utility is the highest which implies a linear preference function in the probabilities of the outcomes occurring. This choice rule can be readily applied to the brand selection problem. Note, however, that the expected utility becomes a random variable in this case because of the random utility assumption.

If $P_i(j \in M)$ denotes the likelihood of individual consumer i judging j to belong to choice set M, expected utility maximization implies that the probability the consumer will select brand j out of choice set M equals

$$P_{ij} = \text{Probability}\left[P_i(j \in M).\ U_{ij} > P_i(k \in M).\ U_{ik} \text{ for } k \in J,\ k \neq j\right]$$

where J denotes the universal set of all brands and U_{ij} is defined as in (1). The inequality in brackets captures the relative magnitude of the expected utilities which are the random outcomes of each "brand lottery". Given the i.i.d. assumption on the random component of the utilities, this choice probability equals

$$P_{ij} = P_i\,(j \in M) \int_{x=-\infty}^{+\infty} g(\mu_{ij} = x) \prod_{\substack{k \in J \\ k \neq j}} P_i\,(k \in M) \int_{y=-\infty}^{V_{ij} - V_{ik} + x} g(\mu_{ik}=y).dy.dx$$

which, when the μ_{ij}'s in (1) are distributed according to a Type - 1 Extreme Value distribution results in

$$P_{ij} = \frac{[\exp(V_{ij})].\ P_i(j \in M)}{\sum_{k \in J} [\exp(V_{ik})].\ P_i(k \in M)} \tag{2}$$

Expression (2) describes the choice probabilities of the basic choice model derived in this paper. These probabilities were derived recognizing the characteristics of the goal-derived categorization process which gives rise to choice set M and their impact on the rational brand selection task.

Although the behavioral premises are different, the choice probabilities expressed in (2) are structurally familiar to and related to many known discrete choice models developed and used in the brand selection literature. The probability expressions are entirely compatible with a Luce choice model formulation (Luce 1959). Furthermore, Tversky's Elimination-by-Aspects (EBA) model (Tversky 1972) can be expressed in a form similar to (2) with utility functions defined in terms of product features (attributes) and the inclusion probabilities $P_i(j \in M)$ representing a measure of distinctiveness (Meyer and Kahn 1991). The traditional multinomial logit model (McFadden 1974) and the nested logit model are special cases of the choice probability model in (2). Model (2) reduces to multinomial logit model when $P_i(j \in M) = 1$ for all $j \in J$. The nested logit model is based on the assumption that choice set M is known or deterministically predictable; hence, $P_i(j \in M) = 1$ if $j \in M$ and $P_i(j \in M) = 0$ if $j \notin M$. The derived choice probabilities are also structurally identical to some generalized logit models (see Batsell 1981, Huber and Sewall

1982, Meyer and Eagle 1982, Meyer and Cooper 1988) and the incremental multinomial logit model (Ben-Akiva and Lerman 1985, p. 114). To complete the choice model, an operationalization of $P_i(j \in M)$ is discussed next which views choice set generation as a rational, goal-derived categorization process.

OPERATIONALIZATION OF CHOICE SET GENERATION PROCESS

As the evaluative constructs of attitude, ideal, and attribute structure have been strongly linked to typicality at the brand level, it is imperative that they are captured in the formalization and measurement of $P_i(j \in M)$. Although motivated by a different rationale, the choice literature contains two approaches which could be considered for measuring $P_i(j \in M)$ (Fotheringham 1988). First, $P_i(j \in M)$ could be defined in attribute space as a measure of brand j's dissimilarity to the other brands. The rationale for this approach is that the likelihood of inclusion in the choice set is affected (either positively or negatively) by the degree to which a brand possesses distinctive features. Several formulations on measuring dissimilarity have been suggested in the literature (see, e.g., Batsell 1981, Meyer and Eagle 1982).

Second, $P_i(j \in M)$ could be defined in a perceptual space as a measure of distance (i.e., spatial proximity) to other brands. The rationale is basically the same as in the previous approach but with the recognition that choice set definition could imply spatial evaluations. Several approaches on measuring spatial distances have been suggested in the psychometric and marketing research literature. This second approach is certainly more in line with the literature on categorization and cognitive representation relied upon here. Nevertheless, it still does not recognize the full spectrum of determinants of typicality. Attitude, familiarity, and goal (ideal) elements are missing. Furthermore, neither approach recognizes the nature of the typicality judgments within the goal-derived categorization process. As discussed above, the judgments are relative to prototypical reference points. Therefore, a more comprehensive measurement approach of $P_i(j \in M)$ which addresses these concerns is needed.

The approach suggested here frames choice set generation as a rational process recognizing that prototypical brands are used as cognitive reference points (Mervis and Rosch 1981). It is characterized by two stages. First, the individual consumer defines a set of brands

which, for a specific consumption goal, contain prototypical brands. Because of the uncertainty about which brand is prototypical, the assumption is made that the individual consumer uses a set of prototypical brands as cognitive reference points in his typicality judgments. That particular set of brands is called the "consideration set". Second, each individual brand in the universal set is compared to these prototypical alternatives in order to judge its typicality given the current consumption goal. To recognize the evaluative determinants of the process as well as the relative nature of the judgments, the two-stage process is operationalized as the outcome of relative typicality evaluations contingent on a specific consumption goal.

Specifically, each brand has a typicality function which measures the brand's typicality given a specific consumption goal. In other words, the function captures the likelihood of the brand satisfying the consumption goal. This typicality function consists of a deterministic component and a random component. In other words, the typicality individual consumer i attaches to brand j for consumption goal ℓ equals

$$T_{ij}^{\ell} = t_{ij}^{\ell} + \tau_{ij}^{\ell} \tag{3}$$

where t_{ij}^{ℓ} denotes the deterministic component of the typicality judgement, and τ_{ij}^{ℓ} denotes the random component. The superscript ℓ indicates that typicality is contingent on consumption goal ℓ. The random component captures the individual's uncertainty in the goal-derived typicality judgment. The deterministic component t_{ij}^{ℓ} contains all determinants which have been identified in the categorization literature as influencing typicality. Some of those determinants could also be influencing the deterministic utility component V_{ij} in (1). Brand preference, for example, which has been identified as a determinant of typicality at the brand level in the consumer behavior literature, is likely to affect V_{ij} as well. In essence, t_{ij}^{ℓ} contains all determinants which affect the conditional evaluation of brand j given consumption goal ℓ, where V_{ij} contains all determinants which affect the unconditional evaluation of brand j. Overlap in determinants does not, however, imply dependence. Indeed, as choice set generation and brand selection play different roles, the random component τ_{ij}^{ℓ} in (3) is assumed to be independent of the random utility component μ_{ij} in (1).

Stage 1: Formation of the Consideration Set

The brands that will be considered prototypical by the individual consumer are identified using a sequential sampling process which balances typicality and evaluation cost. Behavioral support for intuitive cost/benefit calculations in such decision strategies is

provided in Payne (1982), Johnson and Payne (1985), Grether and Wilde (1984), and Huber and Klein (1991). The formalization suggests that a brand will be added (deleted) if the expected maximum likelihood of satisfying the consumption goal contained in (n + 1) brands minus the same expected maximum likelihood contained in n brands exceeds (does not exceed) the additional cost of evaluating the typicality of this additional brand. Specifically, the $(n + 1)^{th}$ brand is added as a prototypical brand if

$$E[\max(n+1)] - E[\max(n)] > d_{n+1} \tag{4}$$

where E is the expectations operator, max (n) denotes the maximum likelihood of satisfying the consumption goal or typicality which could be derived from the n prototypical brands, and d_{n+1} denotes the cost of evaluating the typicality of the $(n+1)^{th}$ brand. Using similar arguments, dropping a brand from the consideration set is a result of the reverse inequality.

The resulting sampling process evolves as follows. We start with a consideration set from t-1 of size n (where n < m, m being the total number of brands in J). All other (m - n) brands are candidates to enter the consideration set at time t; all n brands in the consideration set are candidates to leave at time t. The sequence in which brands will be added or dropped depends on the incremental typicality, or the left side of equation (4). For brands not yet included, this means the incremental typicality to be realized by adding the brand to those already in the set. For those brands currently in the set, this means the incremental typicality over that of the other (n-1) brands in the set. For brands not yet in the consideration set, the sequence of processing moves from the brand with the highest incremental typicality to the one with the lowest. In contrast, for those brands already in the set, the sequence moves from the brand with the lowest incremental typicality to the brand with the highest, since the weakest brands are the most likely to be dropped at an early time. To determine whether an add or a drop decision will be made, the incremental typicality of the brands currently in the set is subtracted from the maximum incremental typicality of all brands, either in or out of the set. These differences are then compared to the incremental typicality of the potential entrants to determine the largest value. If a new brand has the largest value, it is considered for inclusion in the set. If an existing brand has the largest value, it is considered for deletion. The implication of this procedure is that the consideration set can expand and contract over time with prototypical brands entering and tending to stay in.

In order to operationalize the rule in (4), the distribution of the maximum typicality and the decision cost of evaluating the typicality of an additional brand need to be derived.

Assuming that the random typicality components τ^{ℓ}_{ij} in (3) are i.i.d. according to a type-1 Extreme Value distribution, the cumulative density of the maximum given n brands equals (see, e.g. Ben-Akiva and Lerman, 1985)

$$F_{max_i}(x) = \exp\left[-\exp\left(-\left(x - b^{\ell}_i(n)\right)\right)\right] \tag{5}$$

which is a Type-1 Extreme Value distribution with modal value equal to

$$b^{\ell}_i(n) = \ln\left[\sum_{j=1}^{n} \exp\left(t^{\ell}_{ij}\right)\right].$$

Given that the mean of the Type-1 Extreme Value distribution is a constant (Euler's constant) away from the modal value (Johnson and Kotz 1970), the left side of the rule in (4) can be expressed as $E[\max(n+1)] - E[\max(n)] = b^{\ell}_i(n+1) - b^{\ell}_i(n)$. Hence,

$$E[\max(n+1)] - E[\max(n)] = \ln \frac{\sum_{j=1}^{n+1} \exp\left(t^{\ell}_{ij}\right)}{\sum_{j=1}^{n} \exp\left(t^{\ell}_{ij}\right)}$$

which equals the log of the odds ratio of selecting any brand from a set of n versus selecting it from a set of $(n+1)$ brands. Note that the ratio is invariant up to an additive constant just as typicality is determined up to an additive constant.

The evaluation cost d_{n+1} in (4) can be operationalized using the "cost of thinking" framework described in Shugan (1980). This operationalization is not without criticism nor is it the only one suggested in the literature (see Payne 1982). As a comprehensive measure of psychological inspection cost, it does provide a useful starting point. Shugan (1980) postulates that the cost of evaluating different brands against one another is directly proportional to the perceptual complexity in comparing brands, and is inversely related to both the difference in preference between the brands (see also Meyer 1982, p. 106), and the confidence at which the selection must be made. Specifically, Shugan suggests that individual i's potential cost of comparing the typicality of two brands j and k equals

$$f_{ip} = \frac{\mathrm{Var}\left(T^{\ell}_{ij}\right) + \mathrm{Var}\left(T^{\ell}_{ik}\right)}{(1-\alpha)\left[E\left(T^{\ell}_{ij}\right) - E\left(T^{\ell}_{ik}\right)\right]^2}$$

where α denotes the confidence level at which the selection must be made and the obvious restriction that $E(T_{ij}^{\ell}) \neq E(T_{ik}^{\ell})$. Given the random typicality component assumptions made above, it can be shown that

$$f_{ip} = \frac{2\beta}{(1-\alpha)\left[t_{ij}^{\ell} - t_{ik}^{\ell}\right]^2}$$

where β denotes the constant variance of the Type-1 Extreme Value distribution (see Johnson and Kotz 1970, p. 278). One of the implications of the constant variance is that the scale invariance of Shugan's (1980) general results has been lost.

Following the arguments of Shugan (1980) where the cost of evaluation is independent for each brand, the cost of evaluating the typicality of the additional brand can then be expressed as

$$d_{n+1} = \frac{2\beta}{(1-\alpha)} \sum_{j \neq n+1} \frac{1}{\left(t_{ij}^{\ell} - t_{in+1}^{\ell}\right)^2}.$$

Moreover, the decision to add the (n + 1)th brand to the consideration set would depend on the inequality

$$\ln \frac{\sum_{j=1}^{n+1} \exp\left(t_{ij}^{\ell}\right)}{\sum_{j=1}^{n} \exp\left(t_{ij}^{\ell}\right)} > \frac{2\beta}{(1-\alpha)} \sum_{j \neq n+1} \frac{1}{\left(t_{ij}^{\ell} - t_{in+1}^{\ell}\right)^2}. \tag{6}$$

Dropping the (n+1) th brand already belonging to the consideration set would depend on the reverse inequality. Hence, with the add/drop sequence as discussed above, one proceeds until a change in the consideration set occurs (either a drop or an add depending on the inequality); at that point the incremental typicality is recalculated on the basis of the new set and a new sequence is determined. One then proceeds through the new sequence until no changes occur and the composition of the consideration set at time t is defined. A numerical example illustrating the sampling process is discussed in Appendix A.

Note that the right hand side of equation (6), the decision cost, has a number of appealing properties. First, as the factor $(1 - \alpha)$ decreases, decision costs increase. As parameter α can be interpreted as a measure of involvement, this suggests that lower involvement makes

the consideration set expand, which is consistent with Sherif and Hovland's (1964) notion that the latitude of acceptance expands with low levels of involvement (Assael 1984, p. 95). Second, the higher the difference in the deterministic component of typicality between the brands, the lower the decision cost. One could argue that the decision cost is limited because of the ease of discrimination in terms of typicality. Moreover, the directional properties of the decision cost components are theoretically appealing. Adaptations of (6) to ensure scale invariance, to account for inertia and learning in sequential evaluation, and to prevent infinite decision costs for brands with identical mean typicalities are discussed and illustrated in Vanhonacker (1993).

Stage 2: Measurement of $P_i(j \in M)$

When a brand enters the consideration set C, it will raise the maximum typicality contained in the set. That maximum represents in a composite sense how well the consumer could do typicality wise at this time. It embeds a composite score of prototypicality. It is argued here that this maximum forms the cognitive reference point against which the consumer will evaluate the typicality of other brands in the universal set. In other words, the likelihood of being in the choice set is defined as the probability that the perceived typicality of a brand is larger than the maximum typicality contained in the consideration set. Specifically,

$$P_i(j \in M) = \text{Prob.}\left(T_{ij}^{f} > b_i^{f}\right)$$

and, hence,

$$P_i(j \in M) = \text{Prob.}\left(\tau_{ij}^{f} > b_i^{f} - t_{ij}^{f}\right)$$

where b_i^{f} denotes the modal value of the maximum density in (5). The modal value is used on the basis of (a) the skewed nature of the underlying distribution, (b) analytic simplicity, and (c) the fact that the mean is a constant away from the modal value. Furthermore, it makes intuitive sense to use the most likely value of the maximum as the cognitive reference point. Moreover,

$$P_i(j \in M) = 1 - \exp\left[-\exp\left[-\left(b_i^{f} - t_{ij}^{f}\right)\right]\right] \tag{7}$$

with $b_i^{f} = \ln\left[\sum_{j \in C} \exp\left(t_{ij}^{f}\right)\right]$ as defined above.

This probability of belonging to the choice set has some interesting properties. The derivative of P_i ($j \in M$) in (7) with respect to b_i^ℓ is clearly negative. Accordingly, anything which increases the modal value of the maximum will reduce the likelihood of belonging to the choice set for later brands. Analytic results in Appendix B show that this likelihood increases with: (a) a decrease in the number of prototypical brands already contained in consideration set, (b) a decrease in the mean and variance of the typicality judgments of the brands in the consideration set; and (c) an increase in the uncertainty or noise in the typicality judgments. Some interesting strategic conclusions could be drawn from these results. For example, condition (c) implies that a possible strategy for a new brand entering an established market could be to create noise or even question the brand's typicality judgments made by consumers. In this instance, the likelihood of the new entrant being in the choice set would be enhanced.

CHARACTERISTICS AND IMPLICATIONS OF THE CHOICE MODEL

Integrating the operationalization of the choice set generation process into the choice model shown in (2) implies that the probability that consumer i will choose brand j becomes

$$P_{ij}^{\ell} = \frac{[\exp(V_{ij})]\left[1 - \exp\left(-\exp\left(-\left(b_i^{\ell} - t_{ij}^{\ell}\right)\right)\right)\right]}{\sum_{k \in J} [\exp(V_{ik})]\left[1 - \exp\left(-\exp\left(-\left(b_i^{\ell} - t_{ik}^{\ell}\right)\right)\right)\right]} \tag{8}$$

The superscript ℓ denotes that the probability is conditional on the consumption goal ℓ. Under the assumptions about typicality postulated here, the model is conditional on the consumption goal but it is not limited to that. Indeed, evidence exists that typicality is determined by versatility. Ratneshwar and Shocker (1991) have demonstrated in a goal-derived categorization framework that typicality is strongly related to versatility defined as the number of usages for which a brand is perceived appropriate. Accordingly, a brand which is viewed as being helpful to achieve various consumption or usage goals would be viewed as more typical than one that is viewed as being helpful in achieving one single goal. Note that this would imply that the cognitive reference point (the prototypical item) used in typicality judgments would be constant across consumption or usage goals. Accordingly, neither $P_i(j \in M)$ nor the choice probabilities P_{ij}^{ℓ} would be conditional on the consumption goal.

The basic choice model derived here and shown in (8) has some interesting and intuitively appealing characteristics. First, the model does not contain the unappealing Independence of Irrelevant Alternatives (IIA) property which characterizes many choice models (Currim 1982). This property implies that when a new brand is introduced, it will obtain market share proportionally from all brands regardless of substitutability. In the above model, the new brand will obtain proportionally more market share from the prototypical brands (i.e., the brands belonging to the consideration set) regardless of whether or not it is similar to them. The latter is in line with asymmetric dominance which was shown to exist in Huber, Payne, and Puto (1982).

Second, the choice model also does not contain the unappealing regularity property implied by the traditional logit model (Huber, Payne, and Puto 1982). The regularity property refers to the fact that the choice probability of a brand cannot increase by adding a new brand to the choice set. This property arises in the multinomial logit model because the numerator of the choice probability for the existing brand does not change while the denominator increases. This is clearly not the case in choice model (8) where the numerator (and dominator) could change depending on, for example, whether or not the new brand becomes a prototypical brand (and, hence, enters the consideration set).

Third, the choice probabilities are, relative to those implied by the multinomial logit model, more polarized. The implication of incorporating the choice set generation process as operationalized above is that prototypical brands have higher choice probabilities than they would have in a multinomial logit model, and the other brands have comparatively lower probabilities. Hence, prototypicality enhances probability of choice which is in line with findings in the consumer behavior literature that, at the brand level, typicality is related to the brand's positive evaluation (Loken and Ward 1987, Nedungadi and Hutchinson 1985, and Ward and Loken 1988).

One point which requires some discussion is the specification of the deterministic components of utility and typicality (V_{ij} and t^{ℓ}_{ij}, respectively). As pointed out already, the former captures unconditional brand evaluations which affect brand selection where the latter captures brand evaluations conditional on the consumption goal which affect choice set membership. Some determinants, such as brand preferences, influence choice through both effect channels. Accordingly, individual consumer variables as well as external marketing variables which impact on brand preferences should be specified in both V_{ij} and t^{ℓ}_{ij}. Nedungadi (1990) shows experimentally that their effects are quite different depending on the effect channel. Ben-Akiva and Boccara (1990) refer to the differential influence as the "availability" effect versus the "substitution" effect where the latter captures the influence on

the set membership and the latter captures the influence on brand selection. However, not all predictors impact choice through both effect channels.

Some marketing actions taken by firms are likely to be limited in impact to either choice set generation or brand selection. For example, the store environment at the time of choice is more likely to influence the brand selection decision than the typicality judgment underlying choice set generation. A price cut or another form of in-store promotion for a brand may make it temporarily more attractive than other brands without changing their respective typicalities. Clearly, a two-for-the-price-of-one promotional offer does not make that brand more typical for the specific consumption goal. Hence, pricing as well as various forms of promotion are external marketing factors influencing V_{ij} but not t^{ℓ}_{ij}.

One should keep in mind, however, that when the consumer has limited familiarity with a brand (for example, a new brand just introduced), its price point could be used as the focal point from which information about the brand is inferred and generalized (Rips 1975). In other words, the individual consumer might judge typicality on the basis of how close the brand's price is to the prices of other more familiar and prototypical brands. In this instance, the brand's price would be part of the deterministic component of typicality. This expectation formation process very much relates to the pioneering advantage argument in emerging categories put forward by Carpenter and Nakamoto (1989).

Other marketing actions taken by firms are more likely to influence the individual's choice set generation process. Advertising campaigns, certainly when they suggest specific usage occasions for the brands being advertised, are influencing the typicality judgment and should be specified as affecting t^{ℓ}_{ij} in (3). Note that, apart from the content of the advertisement, frequency of exposure might itself enforce typicality or atypicality.

From a managerial perspective, it is important both theoretically and empirically to clarify the impact of external factors. As is evident from the choice model derived here and shown in (8), the overall impact of external marketing actions on the individual's choice is different depending on whether the effect is channeled through the choice set generation process or the choice process. Hence, the normative implications are conditional on the effect channel.

EMPIRICAL ILLUSTRATION

Two empirical illustrations for different product categories are discussed briefly. In the first illustration, typical supermarket scanner data are augmented with survey-based data from

panelists to enable a test of the suggested brand choice model. In the second illustration, the model is estimated on typical scanner purchase data. Focus there is on how the model can be estimated with such commonly available data and how it performs (in terms of fit and insight) relative to the multinomial logit model in this less than ideal data scenario.

In an extensive and joint research effort with an advertising agency conglomerate, A.C. Nielsen - France has been conducting periodic (so far, bi-annual) surveys of its SCAN 7000 panel members. The primary motivation of this effort has been to create a truly single-source data base by complementing traditional scanner panel purchase data with attitude and media exposure data for the same group of consumers. Although still in the development stage, part of the data for the initial 65 weeks of the research experiment covering a single product category were available for analysis.

In contrast to the typical supermarket scanner data which has been used in a number of applications, two additional pieces of information were available for this study. First, telephone survey-based information on aided brand awareness of the scanner panelists was measured. Given evidence in the consumer behavior literature that prototypical items are more salient (Troye 1984, Nedungadi and Hutchinson 1985, Nedungadi 1990), this external information was used to identify the consideration set members and, as such, enabled a full demonstration of the brand choice model developed above. Second, individual-level advertising intensity measures were available for television, radio, and print advertising campaigns. This proprietary measure combines survey-based media habit data with the actual media plans of all the brands in the category, and is derived from the likelihoods the panelist was exposed to each ad insertion. Recognizing learning and forgetting patterns at the individual level, the measure incorporates dynamic carryover effects. For this study, the measure was available on a daily basis (i.e., aggregated across, for example, TV advertising spots throughout the day).

The product category is a frequently-purchased food item used in France almost exclusively as a bread spread. For most households, it is considered a staple. Survey usage data support the assumption that purchases are motivated by a single and identical consumption goal which enabled us to estimate the model across all purchases (i.e., we do not need to separate out purchases motivated by different consumption goals). The category consists of 13 brands, where brand was defined as a combination of brand name and whether the product was of premium quality, regular, or low calorie. The latter product segmentation is known to the consumer, as the particular product type is commonly added as a descriptor to the brand name on the label of the product. Not all brands are available in these three product types.

The scanner panel purchase records of panelists who had made at least three purchases over the 65 week period were used. In total, 32700 purchase acts were incorporated in the estimation. Prior analysis did not reveal any seasonality in the category.

Besides brand constants, 15 predictor variables were specified in the model. Three promotional variables, measured as 0-1 indicators, were incorporated. The first one measured the presence of special in-store displays or placement (e.g., end-of-aisle); the second one measures the availability of a price promotion which could not be cashed in immediately (e.g., coupons); the third one measures the availability of immediately-available price promotions (e.g., cents off). Besides these three promotional variables, actual shelf space occupied by the brand at the time of purchase (measured in cm) was incorporated as a predictor variable. Price was incorporated as relative price, operationalized as the brand's paid price divided by the weighted average of the prices of all brands. The individual-level market (purchase) shares were used as weights so as to capture the likely strength of the price signal to the customer. A loyalty variable operationalized as in Guadagni and Little (1983) was also specified. Advertising intensity was incorporated as a direct brand effect for those brands which had relied extensively on the particular medium in their media plan.

The brand constants were incorporated in the deterministic component of typicality (i.e., t^{ℓ}_{ij} in (3)). As such, they capture the intrinsic prototypicality of the brands for the single consumption goal. As the Guadagni and Little (1983) loyalty variable is an exponentially weighted average of past purchases it captures brand familiarity which, as discussed above, has been linked to typicality in consumer behavior. Accordingly, the loyalty variable was incorporated in the typicality function. Although the contents of the ads throughout the 65 week period were not available, the messages are quite stable in this mature category, and usually incorporate references to the single, dominant usage occasion. Furthermore, as the intensity measures are sensitive to repeated exposures and as such could be interpreted by the consumer as signalling prototypicality, these measures were also integrated as predictors in the typicality function. As amount of shelf space given to a particular brand could be interpreted by the consumer as a measure of its typicality in the product category, shelf space was also integrated as a predictor of typicality. Following the discussion above, the remaining predictors (i.e., promotion and relative price) were specified in the deterministic component of utility.

An overview of the estimation results is provided in Table 1 with the predictor variables listed by specified effect channel. Because of space limitations, the brand constants are not shown. The table contains the aggregate parameter estimates as well as segment-based

Table 1

Estimation Results[a]

Predictor Variables	Across all Buyers	Average Interpurchase Time Segmentation		
		Segment 1 (≤ 18 days)	Segment 2 (between 18 and 36 days	Segment 3 (over 36 days)
1. Utility Functions				
Promotion 1	0.223		0.252	0.237
Promotion 2				1.094
Promotion 3	0.065	0.077		0.079
Relative Price	-0.695	-0.614		-1.318
2. Typicality Functions				
Loyalty	4.043	4.326	4.061	3.437
Shelf Space	0.030	0.040	0.014	0.039
Radio Advertising 1				
Radio Advertising 10				
TV Advertising 1				
TV Advertising 4			2.299	
TV Advertising 6				
TV Advertising 7				
Print Advertising 1				
Print Advertising 6				
Print Advertising 7				
Value of Log Likelihood Function	-5584.718	-1896.393	-1975.808	-1631.632
U^2	0.4650	0.5480	0.4510	0.3850

a Only MLE estimates whose asymptotic t-values were larger than 2 in absolute value are shown; the brand constants were suppressed for space reasons.

parameter estimates. The latter distinguish between three customer segments based on average interpurchase time (i.e., less than or equal to 18 days, between 18 and 36 days, and 36 days or more). The Maximum Likelihood Estimates (MLE) whose asymptotic t-values were larger than two in absolute value are reported together with the value of the log-likelihood function at its maximum and the U^2 goodness-of-fit measure (Hauser 1978). For the latter, the null model assumes that the choice probabilities equal the aggregate market (purchase) shares.

Overall, the fit of the model is good, with the values of the U^2 statistic in the range of values reported in Guadagni and Little (1983). Furthermore, the signs of the various parameters are correct adding to the face validity of the model. With respect to the substantive insights for this illustration, a number of interesting results emerge. In general, the media effects are insignificant. One exception is television advertising for brand 4. That brand ran two one-month campaigns in the second half of the sampling period with an average market-based GRP level of 552. All other brands which relied on TV advertising had campaigns with market-based GRP ratings of 380 or less. Hence, the results indicate that a substantial level of GRPs is needed to impact choice. It is also interesting to note that brand 1 ran an extensive, 4-month radio advertising campaign in the second half of the sampling period. Despite that campaign reaching a market-based GRP rating of 767, no significant impact on choice for that brand was obtained.

With respect to promotional activity, the buyers in this category seem to be quite sensitive and particularly so the more infrequent buyers. Shelf space is significant across all segments. Consistent with previous results with the multinomial logit model, the loyalty is a very significant predictor. As one might expect, the magnitude of its effect is inversely related to purchase frequency. Relative price was also significant with the magnitude of its effect being almost twice as large on infrequent buyers than on frequent buyers. Moreover, the results have intuitive appeal. More importantly, the significant parameter estimates for the typicality predictors support the separate modeling of choice set generation.

A second illustration consists of an example with only scanner purchase data. Given the nature of the data, a couple of caveats might prevent the full power of the model being demonstrated. Nevertheless, this type of data remains commonly available and the question arises how one could parameterize the choice model with such data. More importantly, it provides a strong testing ground for its empirical performance relative to the multinomial logit model.

As the model development suggests, the conditionality of the choice probabilities would require that purchases motivated by different consumption goals are separated out and that a choice model is estimated across purchases for each consumption goal. If all purchases are motived by the same goal (as in the first example just discussed), no particular problem arises as one can readily estimate one choice model across all purchases. But in general, different consumption goals might have been present. As scanner purchase records do not identify the goal for any purchase, we could not separate out the purchases. If we do estimate a single model, we could argue either that average effects across consumption goals are obtained or that versatility determines typicality (for which there is support in the consumer behavior literature as discussed above; see Ratneshwar and Shocker 1991) under which the conditionality of the model would disappear.

A second caveat is the lack of external information on consideration set membership such as the awareness data used above. However, if familiarity and experience are key determinants of prototypicality (for evidence, see Zajonic 1968), consideration could be approximated by purchase set which can be readily determined from the historical purchase records in the scanner data. This is precisely how these caveats were addressed in the next example.

Scanner panel data for a premium beverage category were available from A.C. Nielsen's SCAN 7000 panel in France. Purchase records were available for 829 families who made at least three purchases over a one year period. Because of significant differences in sales levels between the summer and winter months, a distinction was made between high versus low season. Of the total of 6,617 purchases, 4,137 occurred in the high season with the remaining 2,480 in the low season. As in the previous example, the families were segmented based on average interpurchase time. Families with an average interpurchase time of 12 days or less were classified as frequent buyers. Those having a larger average interpurchase time were classified as infrequent buyers. Moreover, the empirical analysis covers two distinct segments over two distinct time periods which captures the extent to which heterogeneity was considered in this example.

The product category has 17 alternatives which were defined as brandname and package size combinations. Three main but distinct brandnames exist (with a fourth covering all other) together with five different package sizes (with a sixth covering all other). Not all brands are available in all sizes. Besides the prices for all alternatives, information on different consumer promotions was also available.

Seven predictor variables were specified in addition to the brand constants. Loyalty, relative price, and three promotional variables were operationalized exactly as in the previous illustration. To capture potential cross price effects, we added two price variables: "name" price and "form" price. These variables were defined as the average price of all other alternatives with either the same brandname or the same package size relative to the weighted average of all prices. These variables will allow us to identify possible brand-primary or form-primary tendencies in price-induced brand switching. As before, the brand constants and the loyalty variable were specified in the typicality function.

The empirical results are summarized in Table 2. The table shows by segment and season the MLE estimates for the traditional multinomial logit model and for the choice model developed above. Estimates with an asymptotic t-value larger than 2 in absolute value are tabulated together with the value of the log-likelihood function and the U^2 goodness-of-fit measure (Hauser 1978). For the latter measure, the null model assumes that the choice probabilities equal the aggregate (purchase) market shares. Again, the brand specific constants are suppressed for space reasons.

Overall, the quality of fit is very good. As one might expect, the model does better for frequent buyers than for infrequent buyers. For frequent buyers, the U^2 is at least 0.64. Guadagni and Little (1983) do not find fits for the multinomial logit better than 0.48. Of particular interest here is the relative improvement of fit for the new model over the multinomial logit model. Note that for corresponding samples, the number of parameters in both models is identical. Indeed, both models are identical with respect to specified predictor variables, but differ in the effect channel of those variables on brand choice. The results in Table 2 indicate that the new model developed here fits better than the multinomial logit model but, as one might expect, the absolute difference in the values of the fit statistic is not large in some cases. As the distribution of the U^2-statistic is unknown, no test statistic exists to judge whether the differences are indeed statistically significant. We pursued a more computer intensive method to clarify the degree of improvement in fit.

For the purchase records in the frequent buyers/low season sample, we selected with a random starting point 50 samples of 1,000 purchases. For each of these subsamples we estimated the multinomial logit model and model (8) specified as above. Across the samples, we obtained an empirical distribution of the U^2 fit statistic. On a one-by-one basis, model (8) always fitted better than the multinomial logit model. For the multinomial logit model, the mean U^2 value is 0.6423 where for the new model it is 0.6460. Assuming normality, these means are significantly different at the 0.025 level (one-sided test). For every U^2 value, the distribution of the fit for model (8) is approximately 12% more to the

Table 2

Comparative Estimation Results[a]

Predictor Variables	Low Season				High Season			
	Frequent Buyers [b]		Infrequent Buyers		Frequent Buyers		Infrequent Buyers	
	Multinominal Logit	New Model (8)	Multinominal Logit	New Model (8)	Multinominal Logit	New Model (8)	Multinominal Logit	New Model (8)
1. Utility Function								
Promotion 1	1.454	1.458	2.967	3.155	1.206	1.222	0.899	0.908
Promotion 2								
Promotion 3	-2.286	-2.2254						
Relative Price								
Name Price				5.420				
Form Price								
2. Typicality Function								
Loyalty	4.525	5.048	3.606	3.251	4.482	4.977	3.889	4.369
Value of Log likelihood function	-976.062	-970.687	-157.439	-148.052	-1 654.708	-1 644.732	-426.773	-425.685
U^2	0.6569	0.6588	0.4529	0.4856	0.6348	0.6420	0.3882	0.3897

a Only MLE estimates whose asymptotic t-values were larger than 2 in absolute value are shown; the brand constants were suppressed for space reasons.

b Average interpurchase time ≤ 12 days.

right. Hence, the results do indicate a statistically significant improvement in fit in this instance.

Substantively, the MLE parameter estimates reported in Table 2 have face validity and provide some interesting insights. The loyalty effect is in general very strong. Loyalty is higher among frequent buyers and more so in the low season. Special display promotions are significant particularly for infrequent buyers in the low season. Immediate price deals are generally not significant, except for frequent buyers in the low season. The instantaneous effect is, however, negative which is somewhat counter-intuitive. Price is generally not an issue, except for infrequent buyers in the low season; there we find a significant positive "name" price parameter indicating that price changes are carried over in the same direction to all alternatives under the same umbrella brandname.

Although the parameter estimates capturing these effects are similar in value for both models, the managerial implications are quite different as a result of the underlying behavioral processes and difference in effect channel as discussed above. The normative implications are indeed very different. As pointed out above, relative to the multinomial logit model, the choice probabilities are more polarized in the new choice model. As these probabilities influence the elasticities to the various external market factors specified, they influence optimal marketing expenditures from a manufacturer's perspective. The behavioral rationale of the model in (8) and its implicit separation between choice set generation and choice enables potentially more efficient marketing spending relative to the multinomial logit model. Consider, for example, the pricing decision. For simplicity, we will take distance from the illustration and assume that prices are specified in the deterministic utility component V_{ij} in (1) in the logarithm of their absolute values. This simple specification immediately illustrates that equality in the price parameter in the multinomial logit model and choice model (8) leads to different short-term price elasticities (and, hence, optimal pricing decisions). Given the logarithmic specification, we can rely on Allenby and Rossi's (1991) results for the multinomial logit model. As shown there, the short-term price elasticity for any brand equals the price parameter (specified to be positive) times the choice share probability for the brand minus one. In model (8), exactly the same result is obtained irrespective of whether the brand is prototypical (i.e., belongs to the consideration set) or not. If the price parameters are identical, the elasticities will still be different because of differences in the second component (as a result of the polarization in shares). In the model developed here and shown in (8), if the brand is prototypical, its choice share will be larger and, hence, the price elasticity will be smaller relative to the one derived from the multinomial logit model. The opposite is true when the brand is not prototypical. Hence, although the estimated price parameters can be identical, the model

developed above recognizes heterogeneity in price response depending on whether the brand is prototypical or not. This heterogeneity, and the behavioral support underlying it and captured in model (8), is ignored in the multinomial logit model.

CONCLUSION AND FUTURE RESEARCH

In this paper, an individual-level choice model was developed which integrates a choice set formation process into the brand selection task. Choice set formation is conceptualized as a goal-derived categorization process and is operationalized as a rational, two-stage process based on individual-level judgments about the typicality of each brand for a specific consumption goal which motivates the choice. Uncertainty surrounding the latter gives rise to a probabilistic definition of choice set membership and to the framing of brand selection as a decision under uncertainty. Within a random utility framework, the individual consumer is assumed to maximize expected utility which results in a comprehensive choice model whose probabilistic predictions are conditional on the consumption goal.

The modelling framework brings together a number of different areas. The categorization paradigm and the typicality construct find strong support in the cognitive psychology literature. Recently, the consumer behavior literature has further expanded that work primarily looking at typicality and its determinants at the brand level. Most of this work has been conceptual, however, and this paper provides an initial but theoretically motivated attempt to operationalize the concept. Another contribution to the literature on categorization and typicality is the explicit recognition of an element of uncertainty underlying these constructs so far not recognized in the consumer behavior literature.

Choice set generation as goal-derived categorization fits into the current stream of research in consumer behavior which claims that all purchase behavior is aimed to achieve certain goals. This paper provides a formalization of that paradigm within a random utility framework. In contrast to all previous discrete choice models, the choice model developed here recognizes that the individual consumer will attach uncertainty to the alternatives in the choice set. Moreover, the framework integrates conceptual developments in the consumer behavior literature on goal-derived categorization with developments in decision sciences and economics on rational decision making under uncertainty. Despite its theoretical richness, the derived choice model remains parsimonious.

The choice model opens up an interesting avenue for further research, particularly in the direction of pertinent managerial questions such as order-of-entry (or pioneering advantages) in emerging product categories and successful brand extensions. The concept of typicality has featured centrally in some innovative work on pioneering advantages (see, Carpenter and Nakamoto 1989, Carpenter et. al. 1993). The model developed in this research provides an integrative analytic framework to possibly study whether or not these advantages exist and if they do, what gives rise to them. The most important application of the typicality construct in the consumer behavior literature has been its use in explaining and predicting the success of brand extensions (see, e.g., Herr et. al. 1990, Boush and Loken 1991, Dawar and Anderson 1993). Again, the choice model framework developed in this paper provides a new vehicle to study questions in this important area.

Besides further theoretical and conceptual developments linked to the typicality construct, the choice model itself provides a number of challenges. First of all, more extensive and comprehensive empirical work is needed. The illustrations discussed in this paper point to some important issues ranging from data requirements to measurement questions. To the extent that external data beyond typical supermarket scanner data are not available, more work needs to be done on operationalizing typicality and, hence, consideration set membership. The extensive cognitive psychology and consumer behavior literatures on the determinants of typicality can perhaps guide us to better surrogate measures when we are only faced with scanner data.

The implicit separation of choice set generation and brand selection and the discussion of various external marketing factors impacting on either or both points to the need for finer measurement of such variables as, for example, advertising. Simple repetition of an ad can enforce typicality, but so can the advertising copy. Furthermore, the conditional character of the derived choice probabilities raises the issue of whether or not it is possible to separate purchases (and how) or whether or not the versatility of a brand across usage occasions is enough of an argument to drop the conditionality. Perhaps in some product categories, package size (now commonly integrated in the brand definition) signals a particular usage occasion. Needless to say, the augmented complexity of the choice model has magnified the question of heterogeneity in estimating the various parameters. The consumer heterogeneity question has recently been looked at extensively in the brand selection part for the multinomial logit choice model (see, e.g., Fader and Lattin 1993), but in the current model the choice set generation process (and, hence, the typicality construct and its determinants) needs to be looked at in the same light.

As discussed in the paper, more research needs to be done on the operationalization of the choice set generation process. Particularly the measurement of the decision costs can be further refined. Some initial work is already reported in Vanhonacker (1993). The latent character of this process makes external validation very difficult. Nevertheless, to the extent that the operationalization can be perfected, it might be possible to implicitly (iteratively) estimate the consideration set (and, hence, prototypicality) given only the individual purchase histories.

The uncertainty implied in the typicality judgment was assumed to be resolved in the choice task. This implies that brand selection is a decision under uncertainty. Each brand has an uncertainty (or risk) associated with it that it does not help in achieving the consumption goal motivating the choice decision. Accordingly, each brand implies a lottery with random utility outcomes. The binomial probability distribution within each lottery reflects the goal-derived typicality judgment underlying the choice set generation process. In this paper, the choice rule across the lotteries was assumed to be maximum expected utility. However, in the extensive literature on decision under uncertainty and expected utility (see, e.g., Fishburn 1988), numerous examples have shown this to be violated continuously. Accordingly, an interesting avenue for further research is to integrate that work within the choice model framework advocated here. Preference formation across lotteries and their impact on probability distribution adjustments and random utility outcomes (commonly viewed in discrete choice modeling as determining brand preference) need to be investigated further.

Finally, the insights discussed in the conceptual development of the model and the empirical illustrations point to a fruitful avenue for managerially-useful normative work. The separation of choice set generation and brand selection implies a heterogeneity with respect to the consumer's response to external marketing factors. More conceptual and empirical work on either process will enable us to further characterize the impact of those external marketing factors and derive more optimal decision rules for them. The behavioral richness of the model provides potentially useful information for manufacturers on more efficient and effective marketing programs.

REFERENCES

Allenby, G.M. and P.E. Rossi (1991), "There is No Aggregation Bias: Why Macro Logit Models Work, *Journal of Business and Economic Statistics*, a (January), 1-14.

Anderson, J.R. (1991), "The Adaptive Nature of Human Categorization," *Psychological Review*, 98, 409-429.

Assael, H. (1984). *Consumer Behavior and Marketing Action.* (Second Edition). Boston, Mass: Kent Publishing Co.

Barsalou, L. (1983), "Ad Hoc Categories," *Memory and Cognition*, 11 (August), 211-227.

Barsalou, L. (1985), "Ideals, Central Tendency, and Frequency of Instantiation as Determinants of Graded Structure in Categories," *Journal of Experimental Psychology*, 11 (October), 629-649.

Batsell, R.R. (1981), "A Multiattribute Extension of the Luce Model Which Simultaneously Scales Utility and Substitutability," in J. Huber (ed.), *The Effect of Item Similarity on Choice Probabilities*, Fuqua School of Business, Duke University, 4-23.

Ben-Akiva, M. (1977), "Choice Models with Simple Choice Set Generation Processes," Working Paper, Dept. of Civil Engineering, M.I.T., Cambridge, MA.

Ben-Akiva, M. and S.R. Lerman (1985). *Discrete Choice Analysis*. Cambridge, MA: MIT Press.

Ben-Akiva, M. and B. Boccara (1990), "Discrete Choice Models with Latent Choice Sets," Working Paper, Dept. of Civil Engineering, M.I.T., Cambridge, MA.

Bettman, J.R. (1979). *An Information Processing Theory of Consumer Choice.* Reading, MA: Addison-Wesley.

Bettman, J.R. (1987), "Process of Adoptivity in Decision Making", in M.J. Houson (ed.), *Advances in Consumer Research*, 15, 1-4.

Boccara, B. (1989). *Modelling Choice Set Formation in Discrete Choice Models*. Ph.D. Dissertation, Dept. of Civil Engineering, MIT, Cambridge, MA.

Boush, D. and B. Loken (1991), "A Process-Tracing Study of Brand Extension Evaluation," *Journal of Marketing Research*, 28 (February), 16-28.

Carpenter, G. and K. Nakamoto (1989), "Consumer Preference Formation and Pioneering Advantage," *Journal of Marketing Research*, 24 (August), 285-298.

Carpenter, G., D.R. Lehmann, K. Nakamoto, and S.B. Walchli (1993), "Pioneering Disadvantage: Consumer Response to Differentiated Entry and Defensive Imitation," Working Paper (June), J.L. Kellogg Graduate School of Management, Northwestern University.

Chintagunta, P.K., D.C. Jain, and J.J. Vilcassim (1991), "Investigating Heterogeneity in Brand Preferences in Logit Models for Panel Data," *Journal of Marketing Research*, 28 (November), 417-428.

Corstjens, M.L. and D.A. Gautschi (1983), "Formal Choice Models in Marketing," *Marketing Science*, 2 (Winter), 19-56.

Currim; I.S. (1982), "Predictive Testing of Consumer Choice Models Not Subject to Independence of Irrelevant Alternatives," *Journal of Marketing Research*, 19 (May), 208-222.

Dawar, N. and P. Anderson (1993), "The Effects of Order and Direction on Multiple Brand Extension," Working Paper (June), INSEAD, Fontainebleau, France.

Fader, P.S., J.M. Lattin, and J.D.C. Little (1992), "Estimating Nonlinear Parameters in the Multinomial Logit Model," *Marketing Science*, 11, 372-385.

Fader, P.S. and J.M. Lattin (1993), "Accounting for Heterogeneity and Nonstationarity in a Cross-Sectional Model of Consumer Purchase Behavior," *Marketing Science*, forthcoming.

Fishburn, P.C. (1988), *Nonlinear Preference and Utility Theory*. Baltimore, Maryland: John Hopkins University Press.

Fotheringham, A.S. (1988), "Note: Consumer Store Choice and Choice Set Definition", *Marketing Science*, 7, 299-311.

Fried, L.S. and K.J. Holyoak (1984), "Induction of Category Distributions: A Framework for Classification Learning," *Journal of Experimental Psychology*, 10, 234-257.

Gensch, D.H. and W.W. Recker (1979), "The Multinomial, Multiattribute Logit Choice Model", *Journal of Marketing Research*, 16, 1124-132.

Grether, D. and L. Wilde (1984), "An Analysis of Conjunctive Choice: Theory and Experiments," *Journal of Consumer Research*, 10 (March), 373-385.

Guadagni, P.M. and J.D.C. Little (1983), "A Logit Model of Brand Choice", *Marketing Science*, 2, 203-238.

Gupta, S. (1988), "Impact of Sales Promotions on When, What, and How Much to Buy", *Journal of Marketing Research*, 25, 342-355.

Hauser, J.R. (1978), "Testing the Accuracy, Usefulness, and Significance of Probabilistic Choice Models: An Information Theoretic Approach," *Operations Research*, 26 (May), 406-421.

Hauser, J.R. and B. Wernerfelt (1989), "The Competitive Implications of Relevant-Set/Response Analysis", *Journal of Marketing Research*, 26, 391-405.

Hauser, J.R. and B. Wernerfelt (1990), "An Evaluation Cost Model of Consideration Sets", *Journal of Consumer Research*, 16, 393-408.

Herr, P., P. Farquhar and R. Fazio (1990), "Extending Brand Equity to New Categories," Working Paper, University of Indiana.

Huber, J. and M. Sewall (1982), "Market Structure From Indices of Competitive Substitutability," in A. Shocker and R. Srivastava (eds.), *Analytic Approaches to Product and Market Planning*, Marketing Science Institute, 7-16.

Huber, J., J.W. Payne, and Ch. Puto (1982), "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis," *Journal of Consumer Research*, 9 (June), 90-98.

Huber, J. and N.M. Klein (1991), "Adapting Cutoffs to the Choice Environment: The Effects of Attribute Correlation and Reliability," *Journal of Consumer Research*, 18 (December), 346-357.

Johnson, M. and C. Fornell (1987), "The Nature and the Methodological Implications of the Cognitive Representation of Products," *Journal of Consumer Research*, 14 (September), 214-228.

Johnson, E.J. and R.J. Meyer (1984), "Compensatory Choice Models of Non-Compensatory Choice Processes: The Effect of Varying Context," *Journal of Consumer Research*, 11 (June), 528-541.

Johnson, E.J. and J.W. Payne (1985), "Effort and Accuracy in Choice," *Management Science*, 30, 1213-1231.

Johnson, N.L. and S. Kotz (1970). *Continuous Univariate Distributors* - 1. New-. York, NY: J. Wiley & Sons.

Lattin, J.M. and R.E. Bucklin (1989), "Reference Effects of Price and Promotion on Brand Choice Behavior", *Journal of Marketing Research*, 26, 299-310.

Loken, B. and J. Ward (1987), "Measures of the Attribute Structure Underlying Product Typicality," in M. Wallendorf and P. Anderson (eds.), *Advances in Consumer Research*, 14, 22-26.

Loken, B. and J. Ward (1990), "Alternative Approaches to Understanding the Determinants of Typicality," *Journal of Consumer Research*, 17 (September), 111-126.

Luce, R.D. (1959). *Individual Choice Behavior*. New York, NY: J. Wiley & Sons.

Machina, M.J. (1987), "Choice Under Uncertainty: Problems Solved and Unsolved," *Economic Perspectives*, 1 (Summer);, 121-154.

McFadden, D. (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka (ed.), *Frontiers of Econometrics*, Academic Press.

Mervis, C. and E. Rosch (1981), "Categorization of Natural Objects," *Annual Review of Psychology*, 32, 89-115.

Meyer, R.J. (1982), "A Descriptive Model of Consumer Information Search Behavior," *Marketing Science*, 1 (Winter), 93-121.

Meyer, .R.J. and T.C. Eagle (1982), "Context-Induced Parameter Instability in a Disaggregate-Stochastic Model of Store choice," *Journal of Marketing Research*, 19 (February), 62-71.

Meyer, R.J. and L.G. Cooper (1988), "A Longitudenal Choice Analysis of Consumer Response to a Product Innovation," in R.G. Colledge and H. Timmermans (eds.), *Behavioral Modelling in Geography and Planning*, London, UK: Croom-Helm, 424-450.

Meyer, R.J. and B.E. Kahn (1991), "Probabilistic Models of Consumer Choice Behavioral," in T.S. Robertson and H.H. Kassarjian (eds.), *Handbook of Consumer Behavior*, Englewood Cliffs, NJ: Prentice-Hall Inc., 85-123.

Nedungadi, P. (1987). *Formation and Use of a Consideration Set: Implications for Marketing and Research on Consumer Choice*. Ph.D. Dissertation, University of Florida, Gainesville, FL.

Nedungadi, P. (1990), "Recall and Consumer Consideration Sets: Influencing Choice without Altering Brand Evaluations, *Journal of Consumer Research*, 17 (December), 263-276.

Nedungadi, P. and J. Hutchinson (1985), "The Prototypicality of Brands: Relationships with Brand Awareness, Preference and Usage," in E. Hirschman and M. Holbrook (eds.), *Advances in Consumer Research*, 12, 498-503.

Nedungadi, P. and V. Kanetkar (1992), "Incorporating Consideration Sets into Models of Brand Choice," in *Advances in Consumer Research*, 19, 251-254.

Payne, J. (1982), "Contingent Decision Behavior", *Psychological Bulletin*, 92, 382-402.

Pitschke, S.B. (1980). *Choice Set Formulation for Discrete Choice Models*. M.S. Thesis, Dept. of Civil Engineering, M.I.T., Cambridge, M.A.

Ratneshwar, S. and A. Shocker (1991), "Substitution-In-Use and the Role of the Usage Context in Product Category Structures," *Journal of Marketing Research*, 28 (August), 281-295.

Richardson, A. (1982), "Search Models and Choice Set Generation," *Transportation Research A*, 16, 403-419.

Rips, L. (1975), "Inductive Judgments about Natural Categories," *Journal of Verbal Learning and Verbal Behavior*, 14, 665-681.

Roberts, J.H. (1989), "A Grounded Model of Consideration Set Size and Composition," *Advances in Consumer Research*, 16, 749-757.

Roberts, J.H. and James M. Lattin (1991), "Development and Testing of a Model of Consideration Set Composition," *Journal of Marketing Research*, 28 (November), 429-440.

Rosch, E. (1973), "On the Internal Structure of Perceptual and Semantic Categories." in T. Moore (ed.), *Cognitive Development and the Acquistion of Language*, New York, NY: Academic Press, 111-144.

Rosch, E. (1975), "Cognitive Representations of Semantic Categories," *Journal of Experimental Psychology*, 104, 192-233.

Rosch, E. and C. Mervis (1975), "Family Resemblances: Studies on the Internal Structure of Categories," *Cognitive Psychology*, 7, 573-605.

Schoemaker, P.J.H. (1982), "The Expected Utility Model: Its Variants, Purposes, Evidence, and Limitations," *Journal of Economic Literature*, 20 (June), 529-563.

Sherif, M. and C.E. Hovland (1964). *Social Judgment.* New Haven, Conn.: Yale University Press.

Shocker, A.D., M. Ben-Akiva, B. Boccara, and P. Nedungadi (1991), "Consideration Set Influences on Consumer Decision-Making and Choice: Issues, Models, and Suggestions," *Marketing Letters*, 2 (August), 181-198.

Shugan, S.M. (1980), "The Cost of Thinking", *Journal of Consumer Research*, 7, 99-111.

Simon, H. (1955), "A Behavioral Model of Rational Choice," *Quarterly Journal of Economics*, 69, 99-118.

Smith, D. and R. Melara (1990), "Aesthetic Preference and Syntactic Prototypicality in Music: Tis the Gift to be Simple," *Cognition*, 34, 279-298.

Snyder, C. and H. Fromkin (1979), *Uniqueness: The Human Pursuit of Difference.* New York, NY: Plenum Press.

Stopher, P. (1980), "Captivity and Choice in Travel-Behavior Models," *Transportation Engineering Journal of ASCE*, 106, 427-435.

Sujan, M. (1985), "The Nature and Structure of Consumer Product Categories," *Working Series in Marketing Research*, number 156, Penn State University.

Swait, J.D. (1984). *Probabilistic Choice Set Generation in Transportation Demand Models.* Ph.D. Dissertation, Dept. of Civil Engineering, M.I.T., Cambridge, MA.

Swait, J.D. and M. Ben-Akiva (1986), "Analysis of the Effects of Captivity on Travel Time and Cost Elasticities," Proceedings of the 1985 International Conference on Travel Behavior, April 16-19, 1985, Noordwijk, The Netherlands, 119-134.

Swait, J.D. and M. Ben-Akiva (1987), "Incorporating Random Constraints in Discrete Models of Choice Set Generation, " *Transportation Research B*, 21B, 91-102.

Troye, S. (1984), "Evoked Set Formation as a Categorization Process," in T. Kinnear (ed.), *Advances in Consumer Research*, 11, 180-186.

Tversky, A. (1972), "Elimination By Aspects: A Theory of Choice," *Psychological Review*, 79, 281-299.

Vanhonacker, W.R. (1993), "The Dynamics of the Consideration Set Formation Process: A Rational Modeling Perspective and Some Numerical Results," Working Paper, INSEAD, Fontainebleau, France.

Ward, J. and B. Loken (1986), "The Quintessential Snack Food: Management of Product Prototypes," in R. Lutz (ed.), *Advances in Consumer Research*, 13, 126-131.

Ward, J. and B. Loken (1988), "The Generality of Typicality Effects on Preference and Comparison: An Exploratory Test," in M. Houston (ed.), *Advances in Consumer Research*, 15, 55-61.

Williams, H. and J. Ortuzar (1982), "Behavioural Theories of Dispersion and the Mis-Specification of Travel Demand Models," *Transportation Research B*, 16B, 167-219.

Wright, P. and F. Barbour (1977), "Phased Decision Strategies: Sequels to Initial Screening," in M.Starr and M. Zeleny (eds.), *Multiple Criteria Decision Making*, Amsterdam, NL: North-Holland, 91-109.

Zajonic, R. (1968), "Attitudenal Effects of Mere Exposure," *Journal of Personality and Social Psychology*, 9, 1-27.

Appendix A: Numerical Example of Consideration Set Formation Process

Assume the deterministic typicality values for 5 brands (m = 5) at time t are,

$t^l(1) = 6$	$t^l(3) = 8$	$t^l(5) = 5$
$t^l(2) = 2$	$t^l(4) = 5$	

Futhermore, the composition of the consideration set at t - 1 is: 1 0 0 1 0, or $\{1, 4\}$ (and, hence, n = 2), and α is set at 0.5. Accordingly, we can compute the incremental typicality (add decisions) and the relative incremental typicality (drop decisions):

	C (I)	exp (t^l(I))	Incremental Typicality	Max - Incremental Typicality
$t^l(1) = 6$	1	403.43	ln (1 + 2.7183) = 1.3133	0.5433
$t^l(2) = 2$	0	7.39	ln (1 + 0.0134) = 0.0133	-
$t^l(3) = 8$	0	2980.95	ln (1 + 5.4018) = 1.8566	-
$t^l(4) = 5$	1	148.41	ln (1 + 0.3679) = 0.3133	1.5433
$t^l(5) = 5$	0	148.41	ln (1 + 0.2689) = 0.2382	-

Hence, the derived sequence of consideration is: 3 (add), 4 (drop), 1 (drop), 1 (drop), 5 (add), 2 (add).

Outcome: 3 cannot be added (incremental typicality < decision cost) (1.8566 < 2.3760)
4 dropped (relative incremental typicality < decision cost) (1.5433 < 6.5798)

With the new consideration set configuration, we have

	C (I)	Incremental Typicality	Max-Incremental Typicality
$t^l(1) = 6$	1	ln (1 + 403.43) = 6.0025	0.0000
$t^l(2) = 2$	0	ln (1 + 0.0183) = 0.0182	-
$t^l(3) = 8$	0	ln (1 + 7.3890) = 2.1269	2.1269
$t^l(4) = 5$	0	ln (1 + 0.3679) = 0.3133	-
$t^l(5) = 5$	0	ln (1 + 0.3679) = 0.3133	-

The sequence of consideration now is 3 (add), 4 (add), 5 (add), 2 (add), 1 (drop).

Outcome: 3 added (incremental typicality = 2.1269 > decision cost = 1.6450).
For the new consideration set configuration, we obtain

	C (I)	Incremental Typicality	Max-Incremental Typicality
$t^i(1) = 6$	1	ln (1 + 0.1353) = 0.1269	2.0000
$t^i(2) = 2$	0	ln (1 + 0.0022) = 0.0022	0.0022
$t^i(3) = 8$	1	ln (1 + 7.3890) = 2.1269	0.0000
$t^i(4) = 5$	0	ln (1 + 0.0439) = 0.0429	0.0429
$t^i(5) = 5$	0	ln (1 + 0.0439) = 0.0429	0.0429

The derived sequence of consideration becomes: 1 (drop), 4 (add), 5 (add), 2 (add), 3 (drop).

Outcome: 1 dropped (incremental typicality = 0.1269 < decision cost = 0.2500).

No brand can be added or dropped subsequently, and the final composition of the consideration set is: 0 0 1 0 0, and hence = {3}.

Appendix B: Likelihood of Being in the Choice Set (P_i ($j \in M$))

The effect of consideration set size, variance and mean of the typicality judgment on likelihood of being included in the choice set can be assessed as follows. By definition,

$$b_i^f = \ln\left[\sum_{j \in C} \exp\left(t_{ij}^f\right)\right].$$

Using a Taylor expansion, b_i^f is approximately equal to

$$b_i^f \cong \ln\left[n + n\,\bar{t}_{ij}^f + (1/2) \sum_{j \in C} \left(t_{ij}^f\right)^2\right] \tag{B-1}$$

given $n\,\bar{t}_i^f = \sum_{j \in C} t_{ij}^f$ where n denotes the number of brands in the consideration set.

Furthermore,

$$\sum_{j \in C} \left(t_{ij}^f\right)^2 = n\,S_i^2 + n\,\bar{t}_i^f$$

with $S_i^2 = \sum_{j \in C} \left(t_{ij}^f - \bar{t}_i^f\right)^2 / n$.

Moreover, expression (B-1) becomes

$$b_i^f \cong \ln\left[1 + \bar{t}_i^f + (1/2)\,\bar{t}_i^f + (1/2)\,S_i^2\right] n.$$

Taking partial derivatives, we obtain

$$\frac{\partial b_i^f}{\partial n} = \frac{1}{n}$$

which is always larger than zero.

Furthermore,

$$\frac{\partial b_i^f}{\partial \bar{t}_i^f} = \frac{(1 + \bar{t}_i^f)}{\left[1 + \bar{t}_i^f + (1/2)\,\bar{t}_i^f + (1/2)\,S_i^2\right]},$$

and

$$\frac{\partial b_i^f}{\partial S_i^2} = \frac{1}{2\left[1 + \bar{t}_i^f + (1/2)\,\bar{t}_i^f + (1/2)\,S_i^2\right]}$$

which are both positive when $\bar{t}_i^f$ is positive.

In sum, all three derivatives are generally positive which together with $\frac{\partial P_i\,(j \in M)}{\partial b_i^f} < 0$ implies that P_i ($j \in M$) decreases as the number of prototypical alternatives, their mean typicality, and the variance in their typicality increases.

The effect of uncertainty (or noise) in the typicality judgments can be assessed by adding an uncertainty parameter to the Type-1 Extreme Value distribution as follows

$$p\,(\tau_{ij}^f) = \Theta_i^{-1} \exp\left[-\frac{\tau_{ij}^f}{\Theta_i} - \exp\left(\frac{\tau_{ij}^f}{\Theta_i}\right)\right]$$

where the variance of the random component τ_{ij}^f is proportional to parameter Θ_i. Hence, larger values of Θ_i would indicate increased uncertainty in typicality. It is straightforward to show that

$$\frac{\partial P_i\,(j \in M)}{\partial \Theta_i} > 0$$

which implies that the likelihood increases with increased uncertainty (or noise).