MODELING CATEGORIZATION, ELIMINATION, AND CONSIDERATION FOR NEW PRODUCT FORECASTING OF CONSUMER DURABLES

by

Glen L. Urban*, John S. Hulland**, Bruce D. Weinberg*

Massachusetts Institute of Technology, Cambridge, MA
Alfred P. Sloan School of Management, W.P. #3206-90

* MIT Sloan School of Management
** University of Western Ontario
ABSTRACT

This paper extends previous models for pre-market forecasting of new durable consumer goods by including the formation of consideration via categorization and elimination processes. After reviewing the relevant behavioral science foundations of categorization, a model and measurement methodology is proposed and applied to the pre-market forecasting of the new Buick Reatta. Empirical data is collected, consideration sub-model parameters are estimated, managerial implications are discussed, and future research needs are identified. The categorization model structure is statistically significant in this application and managerial insight into positioning for effective categorization and buying is generated. Relevance of the categorization extension for other new product models is explored.
ACKNOWLEDGEMENTS

The authors gratefully acknowledge the substantial financial support provided to MIT by the Buick Motor Division of General Motors that allowed us to conduct this research. The second author would also like to acknowledge the financial support of the Social Sciences and Humanities Research Council of Canada for his work.

We appreciate very much the design ideas and research support provided by Deborah Marlino, Frank Kardes, Cheena Srinivaran, Jessamy Tang, and John Dabels.
INTRODUCTION

Many existing consumer product categories can be best characterized as "crowded". An individual consumer making a purchase in one of these product categories faces a large number of possible choice alternatives. For example, the domestic U.S. automobile market is composed of over 300 distinct models; a potential buyer of a personal computer system might face more than 30 choice alternatives; and consumers can choose from among more than thirty different deodorants. While the number of available alternatives in these product categories is large, for any given purchase occasion an individual consumer is unlikely to evaluate all of them. Customers simplify their decision making by eliminating alternatives from consideration. For example, the median number of cars considered by a U.S. consumer is 8.1 (Hauser, Urban, and Roberts 1983)^1

It is important for managers to understand the nature of these restricted consideration sets because the marketing strategy for gaining entry into consumers' consideration sets may be very different from the strategy for maximizing choice once the brand is considered. For example, if 75% of the California, college educated potential auto buyers aged 30 to 45 with income over $35,000 would not consider a Buick, it may not pay to advertise the advantages of Buick relative to Chrysler and Ford. Rather the marketing problem may be how to qualify Buick as relevant choice alternative in consideration sets dominated by Japanese and European cars. This issue is particularly important for new products because if consumers use a consideration heuristic, the

^1. See Hauser and Wernerfelt (1990), p. 394, for a complete tabulation of data on consideration set sizes for packaged goods.
product must be positioned for a categorization that leads to high levels of consideration as well as a large share of choices given that it is considered.

Existing new product forecasting procedures model how consumers gain awareness of a new brand or assume that consumers already employ a particular consideration, evoked, or relevant set of alternatives prior to beginning the analysis. In this paper we extend new durable consumer product forecasting models to explicitly include categorization and consideration. We begin by reviewing relevant behavioral factors that influence the probability of consideration and then incorporate the processes of categorization, elimination of categories, and consideration into an aggregate customer flow model. We propose a measurement and estimation methodology and provide an application in the U.S. automobile market. We end the paper with a discussion of future research needs and potential extensions of the approach employed here.

LITERATURE

Behavioral Foundations

Although consumers are faced with an immense amount of product information, and a relatively large number of alternative brands, most individuals have neither the cognitive abilities nor resources required to effectively process the total set of available information relevant to a particular choice occasion. As a result, many consumers resort to the use of one or more simplifying choice heuristics (see, for example, Bettman 1979, Hogarth 1987, and Payne, Bettman, and Johnson 1988). For example, a growing body of experimental research suggests that individuals use phased decision heuristics when faced
with complex, multi-alternative decision environments (e.g. Payne 1976, Lussier and Olshavsky 1979, Payne, Bettman and Johnson 1988). The first phase of such a strategy involves the use of a noncompensatory decision rule to eliminate most of the alternatives from further consideration. Subsequent phases of this strategy then involve a more detailed, compensatory evaluation of the surviving alternatives, assuming that a single alternative is not chosen in the first phase.

These findings are consistent with a framework originally developed by Howard and Sheth (1969), which suggested that consumers might, when faced with a large number of alternatives, screen the initial set of alternatives down to a much smaller set of relevant brands. They referred to this smaller set of brands as the "evoked set", a set from which all of a consumer's product choices would be made. More recently, this set has been referred to as the consideration set (Alba and Chattopadhyay 1985). Considerable theoretical and empirical work has been conducted over the past two decades to model the size of the consideration (see Hauser and Wernerfelt 1990, and Roberts 1989).

Consideration sets are consistent with finding that consumers organize product knowledge in memory into schemas or categories which structure that knowledge in a meaningful, yet simplified manner. In short, categorization provides consumers with an heuristic that helps simplify their day to day evaluative processes. When a new product is classified by an individual as an example of an existing category, the evaluation associated with that category can be transferred to the new product directly. On the other hand, if the new product is perceived by the individual to not belong to any
existing product category, a detailed evaluation on an attribute by attribute basis becomes necessary (e.g. Sujan 1985). As a result, if a new product is categorized as belonging with other attractive alternatives (e.g. members of the consideration set), it is more likely to receive further attention than if it is categorized as belonging with a set of unattractive products. Incorporating these principles in a new product model has the potential to generate new insights, more accurate forecasts, and more effective marketing strategies.

Aggregate Forecasting Models

Existing aggregate forecasting models have assumed that consumers are already in or move to an awareness state for a new product, or that they have employed a consideration, evoked, or relevant set of alternatives to begin the analysis.

The Bass (1969) product class model and its extensions (see Mahajan, Muller, and Bass 1990), Urban's SPRINTER model (1970), and the brand switching models of Massey, Montgomery, and Morrison (1970) are three well known examples of aggregate models. The Bass and Urban models deal only with a single new product (or product category) and are not capable of incorporating individual consumers' consideration sets. The stochastic models developed by Massey et al. incorporate switching behavior and competitive effects at an aggregate level, but again do not examine individual differences in consideration.

Aggregate choice models that in part deal with consideration sets have been developed (e.g. Silk and Urban 1978, Roberts and Urban 1988, and Guadagni and Little
1983). They do not model customers categorization of acceptable and unacceptable brands, but rather model response to the brands that are in the customers' consideration set.

Conjoint choice models use individual level utility measurements on an attribute by attribute basis. Such models assume a particular set of relevant alternatives and more advanced versions have been used to determine market share through logit or multiattribute analysis (e.g. Green and Srinivasan 1978). While it may be possible to use utility-based choice models that are based on individuals' self-identified consideration sets, researchers to date have commonly preferred to define the relevant sets themselves prior to the utility measurement phase based on hypothetical profiles of attributes which are assumed to be considered by consumers.

To make the commonly employed new product structure of awareness or evoking clearer, consider the new auto forecasting model developed by Urban, Hauser, and Roberts (1990). Figure 1 describes this macro flow model. Customers are defined as being in behavioral process states and as flowing from one state to another as a result of their search for information and marketing activities. For example, customers move from unaware to aware of advertising for the new automobile at some rate per month dependent on the level of advertising. If there are 10 million customers initially unaware and 10% gain awareness in the first month of a new product launch due to advertising, 1 million will be placed in the "aware via ads" box in Figure 1. This represents a conditioning by awareness. If 5% of those who are aware are in the market for a car that month, 50,000 people will flow to the "in market" box. If 30 percent of these visit a
FIGURE 1
Behavioral States Macro-Flow Model for a New Automobile*

*(Adapted from Urban, Hauser, and Roberts 1990)
dealer carrying the new product, 15,000 will be in the "visit dealer" box. Finally, if 75% do not receive word of mouth communication from other customers and 33% of these buy the car, while of the remaining 25% who do receive word of mouth recommendation 50% actually buy the car, the first month sales forecast is 5,625 units (15,000*.75*.33 + 15,000*.25*.5). Other flows are defined to describe forgetting as well as the generation of word of mouth communication and its effect on awareness. The estimated flow parameters are based on market research in a clinic before the new car is introduced. While this model encompasses many of the customer information and decision steps, it does not include the key behavioral factors of consideration through categorization and elimination of alternatives.

In this paper we extend the durable goods new product model shown in Figure 1 to include consideration heuristics, develop a measurement and estimation procedure for categorization and consideration, and empirically apply the new model. The proposed extension has implications for other existing models which condition the new product forecast by awareness, consideration, or evoking of the new brand. We discuss those implications in the final section of this paper.

MODEL STRUCTURE

The extension to include categorization and consideration is conceptually simple. We add new states to represent separate categories of brand alternatives and parameters to define the fraction of brands that are considered in each category. In each category we
FIGURE 2

Expanded Macro-Flow Model, Incorporating Categorization for Advertising Awareness Cell
model the effect of categorization on dealer visits. Figure 2 shows the extended flows for the awareness via ads states of the model for an arbitrary set of nine categories. Similar categorization breakouts are modeled for the aware via word of mouth and aware via both ads and word of mouth states, but are not shown in Figure 2. In each application, the number of categories and their brand alternative composition are determined empirically (see below). The fraction of consumers who are in the market for the overall product class is not expected to vary substantially by category, but the fraction who consider a new car and will visit a dealer is likely to depend on the categorization state. The latter probabilities are expected to be high in categories where all or most of the choice alternatives are considered, and low in categories that contain many eliminated alternatives. Finally, the post visit word of mouth effects and buying probabilities given dealer visit are modelled as in Figure 1.

MEASUREMENT AND ESTIMATION

While the extension to the model structure is straightforward, the increase in the number of states and flows is large and the measurement, aggregation, and estimation procedures involved are more complex. In this section we describe how to estimate the model flows in Figure 2 by measuring individual categorization, modeling probabilities of categorization, elimination, and consideration, and then aggregating the individual categorization data.

To begin the individual categorization measurement, each respondent is given a deck of cards representing the entire set of product alternatives available for his or her
purchase. Each card represents one choice alternative with a photograph on the front and a short list of product attributes on the back. The respondent is asked to identify the subset of products with which he or she is familiar, without referring to the attribute information on the reverse of each card. Once this task has been completed, the unfamiliar choice alternatives are removed. The individual is then given the remaining deck of cards and asked to sort the cards into piles such that any pile will contain only those alternatives that he or she feels belong together based on their similarity (attribute information can be used in this task). Once the piles are formed, the individual is then asked to identify those choice alternatives that he or she would seriously consider for purchase. The resulting subset of alternatives forms that individual's consideration set. The individual is also asked to identify the most typical purchase alternative in each of the piles formed previously so the piles can be given an identifying label and exemplar.

**Categorization, Elimination, and Consideration Probabilities**

Each of the piles formed by an individual in the sorting task can be viewed as a separate sub-category within the overall product category. The categorization process for an individual consumer is measured by the number of piles that he or she forms, and by the composition of those piles. Typically there is a great deal of heterogeneity in both of these measures across individuals, even though the overall product category is common to all.

In order to define the elimination process for an individual, we examine the extent to which the consideration set is represented in the various piles. At one extreme,
an individual might place all of the consideration set choice alternatives in a single pile during the sorting task, even while forming many other piles. At the other extreme, the consideration set elements could be spread out over all of the piles formed. This measure of the span of the consideration set across the formed piles represents the extent to which the individual employs a categorization-based strategy of elimination.

Product consideration is dependent upon the individual's categorization and elimination processes. A new product will be potentially considered only when categorized as belonging to the set of alternatives that are not eliminated outright from further consideration. Even when placed in one of these categories, however, the new brand may or may not subsequently be considered seriously for purchase, because all items in a category may not be considered. (A hypothetical example of one individual's pile formation, categorization, elimination, and consideration processes is provided in the Appendix).

It is possible to describe these processes in a more formal probabilistic manner. Define $n_f$ to be the number of items with which the respondent is familiar, $n_c$ to be the number of items in the individual's consideration set, $n_e$ to be the total number of items that are eliminated from further consideration by being placed in a pile that contains no member of the individual's consideration set, and $n_{e'}$ to be the number of items eliminated from piles that contain at least one considered alternative. (Note: $n_f = n_e + n_{e'} + n_c$). The elimination and categorization processes can be viewed as stochastic in nature. Define $P(E)$ as the probability that an average alternative is eliminated by categorization, $P(\bar{E})$ as the probability that it is not eliminated by
categorization, \( P(C) \) as the probability that it is considered, and \( P(C|\bar{E}) \) as the conditional probability that it is considered given that it is not eliminated initially by categorization. Then:

\[
P(E) = \frac{n_e}{n_f} \tag{1}
\]

\[
P(\bar{E}) = 1 - P(E) = \frac{(n_f - n_e)}{n_f} \tag{2}
\]

\[
P(C|\bar{E}) = \frac{n_c}{(n_f - n_e)} = \frac{n_c}{n_c' + n_c} \tag{3}
\]

\[
P(C) = \left( \frac{n_c}{n_f} \right) = P(C|\bar{E}) P(\bar{E}) \tag{4}
\]

Traditionally, it is \( P(C) \) that has been employed in aggregate flow models whenever individual consideration has been incorporated into the forecasting. However, it should be obvious that equation (4) is formed from the product of equations (2) and (3). Breaking the probability of consideration out into two components permits a deeper examination of the reasons underlying consumer consideration (or rejection). In particular, if consumers tend to view a new product as belonging to an unattractive category and therefore eliminate it outright, the implications for a particular manufacturer's positioning strategy are different than in the case when consumers do not eliminate the product on the basis of categorization alone, but do reject it from serious consideration once they have examined its attributes more extensively. In the former case, a new communications strategy could be used to change consumers' perceptions of the new product, and hence lead them to categorize it as a member of a more attractive category. This change in strategy would be particularly effective if the new product has desirable features that are likely to lead to a high level of consideration, given only that the product is no longer dismissed out of hand. In the latter case, either a product
repositioning strategy or an advertising campaign aimed at changing consumers' attribute weights would have to be employed to change the brand's overall level of consideration.

The preceding results apply to an average item placed randomly in any one of the piles formed by the individual, but it also is useful to examine the elimination and consideration processes on a pile by pile basis. Define the superscript \( j \) to represent the individual categories (or piles) formed in the sort task. Then the total consideration probability for the new product is:

\[
P(C) = \sum_{j=1}^{J} P(C | (E)^j) P(E)^j P(j)
\]

where the expressions from equations (2) and (3) are now determined separately for each category formed, and where \( P(j) \) represents the probability that the new product will be placed in category \( j \).

**Aggregation Issues**

Thus far, the discussion has referred to individual level models of categorization, elimination, and consideration. However, the new product flow model forecasting is done at an aggregate level, so overall categories must be defined to reflect the piles formed by individuals and the average flow proportions calculated for each category. These average flow proportions are estimated by assigning an individual to one and only one of the categories, based on where they place the card for the new brand, and then
averaging the probabilities of considering the new durable product and visiting a retailer across all individuals assigned to a particular category.

It is easiest to think of categories defined at the aggregate level as clusters of choice alternatives. Each choice alternative can be uniquely assigned to one (and only one) of these clusters. The composition of these clusters can in turn be determined by analyzing the information obtained from individuals' card piles. As noted earlier, individuals are asked to form distinct piles such that any two items falling within the same pile are perceived as similar. It is thus possible to form an N x N symmetric similarity matrix, where N is the total number of alternatives in the product class. At an individual level, the matrix element \( (S_{i,k}) \) will equal one if item i and item k were placed in the same pile by that individual, and zero otherwise. If an item i is unfamiliar to that individual, all rows and columns associated with i in the matrix will equal zero.

An aggregate similarity matrix can then be formed by summing these elements across all individuals. At this aggregate level, the off-diagonal matrix elements represent the total number of individuals placing the two items together (i.e. perceiving the two to be similar), while the diagonal elements represent the total number of individuals familiar with the associated alternative. This matrix is converted to a dissimilarity matrix directly, and then input into a hierarchical clustering program. The output of this program identifies both the number of distinct aggregate clusters and the composition of each of these clusters (in terms of the items assigned). The cluster solution obtained using this approach can be tested for reasonableness by using a scree test, or by examining pseudo-F statistics. Alternately, the obtained cluster solution can be
compared to other possible structures determined on an a priori basis by using the statistical testing procedure proposed by Urban, Johnson, and Hauser (1984). This procedure is based on the stated probability of purchase for brands that individual consumers would choose in a category when informed that their first choices are unavailable. This "forced switching" information is then used as an input to a statistical test of the estimated market structure versus a null hypothesis of no structure. A normally distributed Z statistic is generated to test the existence of a categorization structure.

The assignment of the new brand to one of these aggregate clusters is not obvious since an individual is free to place the card for the new brand in any of their individual piles, and can form any number of piles with widely varying compositions. At the aggregate level, each existing alternative is identified with one and only one cluster. It is therefore possible to look at the extent to which the new alternative is perceived by individuals to be similar to (i.e. placed in the same pile with) members of each of these aggregate clusters. This approach leads to the following individual level assignment rule. If the new item is placed with more members of a particular aggregate cluster than with the members of any single other aggregate cluster, then the new item is assigned to that cluster for that individual (see Appendix A for an example of this assignment rule for one individual). When the new item is placed with an equal number of members from two (or more) clusters, the item will be randomly assigned (on an equal weights basis) to one of these clusters. The outcome of the assignment is an estimated probability (P(j)) for the new product being assigned to each of the aggregate categories (j).
The flow model parameters for the proportion of consumers who consider the car and visit a dealer given they categorize the new car in category \( j \) are estimated by averaging the stated probabilities for those individuals assigned to category \( j \).

APPLICATION

An application of the model and measurement methodology has been made to the prelaunch forecasting of the Buick Reatta, a two seat luxury car priced at over \$25,000. The study was done 18 months before introduction. Management was interested in a final sales forecast and targeting and positioning recommendations so that the national launch would maximize sales and profits.

Measurement

Individual measures were collected in a clinic that was run over a three week period in Cincinnatti, Ohio during the summer of 1987. Individuals were invited to participate in the clinic if they were considering a new car for purchase in the next three years, willing to consider purchasing a two seat car, and willing to spend more than \$20,000 for a new automobile. Qualified individuals were then invited to participate in the clinic at their convenience and were offered \$50 as compensation for their time. A total of 434 individuals participated in the study. An average clinic session lasted about two hours.

When a participant arrived at the clinic, he or she was handed a deck of cards representing 69 models of automobiles available domestically within the sporty, luxury car market. This individual was asked to identify the subset of cards with which he or
she was familiar. The average size of this familiarity set across the sample was 56 cars, and it ranged in size across individuals from a minimum of 10 to a maximum of 69. The individual was then asked to sort the set of familiar alternatives into a self-determined number of separate piles based on item similarity. Clinic participants formed an average of 8 distinct piles, with this number ranging from a low of 2 to a high of 27.

Once the content of the formed piles was recorded by the interviewer, the clinic participant was asked to identify the most typical car in each pile formed during the sorting task, the cars that he or she would consider seriously for purchase (i.e. the members of the consideration set), and the cars that represented his or her top three choices for a new automobile purchase. The average size of the consideration set was 7 cars, and the size of this set ranged from a minimum of 1 car to a maximum of 27 cars. The average probability of consideration for this group across all familiar cars was therefore 12.5% (i.e. $P(C) = \frac{7}{56} = 0.125$).

Next respondents saw a concept board which positioned the Buick Reatta as a "new kind of sports car" that combined sporty styling with more comfort and ease of handling. This was the base case positioning for the forecasting effort. Respondents were given a card with the picture of the Reatta and asked to put it either in the pile with those cars they viewed as most similar to it or in a new pile. Additional information about the car was provided by a showroom visit, test drive of a preproduction prototype, and video tapes designed to simulate word of mouth communications. Probabilities of purchase were measured after each information
exposure and respondents were asked to categorize the Reatta after drive, as well as after initial, concept exposure.³

Individual Pile Formation Results

The average clinic participant's consideration set spanned five of the eight piles formed, leaving three piles that were eliminated outright from further consideration. That is, the average individual was unwilling to consider seriously for purchase any of the cars included in these three piles.

By examining the individual pile information we find that our average participants eliminated 46% of their familiar alternatives by placing them in piles that contained no member of their consideration sets (i.e. see equation (1), \( P(E) = 0.46 \) and equation (2), \( P(\bar{E}) = 0.54 \)). All piles that included at least one member of the consideration set survive the category elimination process. However, not all cars in these piles are considered actively for purchase. The average probability of consideration for a car in one of these non-eliminated piles was only 23% (i.e. see equation (3) \( P(C|\bar{E}) = 0.23 \)). The average consideration probability over the individual data was 12.4% (see equation (4), \( P(C) = .23 \times .54 \)).⁴

Aggregate Categorization and Consideration Results

The composition of the piles formed by the 434 clinic participants in the sorting task were used to form a 69 by 69 symmetric similarity matrix as described earlier. This

³ See Urban, Hauser and Roberts (1990) for details of model estimation. We emphasize primarily the new categorization results in this paper.

⁴ As expected, this is similar to the average consideration set size divided by the average number of familiar cars (7/56 = .126).
matrix was converted into a dissimilarity matrix, which was then used as an input into a hierarchical clustering algorithm. Three alternate solutions were obtained, representing 5, 9, and 32 clusters (categories) of automobiles. While the 5 and 9 category solutions yielded similar pseudo-F statistics, the 9 category solution had greater face validity as determined by the relative interpretability of the sets generated in each case. The 32 category solution was less satisfactory from a statistical perspective and difficult to interpret. The composition of the nine category solution is shown in Table 1. In the 5 category solution, categories 1 through 4 of this table were combined, as were categories 5 and 6. The 9 category solution was selected as the best estimate of the aggregate categories. The Urban, Johnson, and Hauser (1984) procedure described earlier was used to statistically test this solution against a hypothesis of no categorization -- all cars compete with each other in one group rather than the nine groups shown in Figure 1. The overall Z statistic was 7.1 and significant at the one percent level; there was significant evidence that the market was divided into categories. All the Z statistics for the categories were significant at least at the five percent level and are shown in Table 1.

The information contained in the piles of individual clinic participants was re-examined using the groups of cars identified in Table 1. Table 2 presents this information by category, aggregated across all individuals (note that the calculations are made prior to examining the location of the Reatta card for each individual). Using category 1 as an example, the cards representing cars in this category (refer to Table 1) showed up in individuals' piles a total of 2015 times. Just over one half of these cards (1061) were located in individuals' piles that contained no members of consideration sets.
### TABLE 1
9 Category Solution

<table>
<thead>
<tr>
<th>Category 1</th>
<th>Category 2</th>
<th>Category 3</th>
<th>Category 4</th>
<th>Category 5</th>
<th>Category 6</th>
<th>Category 7</th>
<th>Category 8</th>
<th>Category 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2.6)*</td>
<td>(2.0)</td>
<td>(1.7)</td>
<td>(2.9)</td>
<td>(3.0)</td>
<td>(2.4)</td>
<td>(3.1)</td>
<td>(3.8)</td>
<td>(4.1)</td>
</tr>
<tr>
<td>Acura Integra</td>
<td>Alfa Rom. GTV</td>
<td>Buick Skyhawk</td>
<td>Honda CRX</td>
<td>Chev Camaro</td>
<td>Chev Corvette</td>
<td>Audi 5000</td>
<td>Buick Electra</td>
<td>Buick Century</td>
</tr>
<tr>
<td>Mazda 626</td>
<td>Maserati Biturbo</td>
<td>Pontiac Sunbird</td>
<td>VW Scirocco</td>
<td>Pontiac Firebird</td>
<td>Porsche 924S</td>
<td>BMW 500</td>
<td>Olds 98</td>
<td>Crys. LeBaron</td>
</tr>
<tr>
<td>VW Jetta</td>
<td>Merkur XR4Ti</td>
<td>Pontiac Grand Am</td>
<td>Honda Prelude</td>
<td>Mazda RX7</td>
<td>Porsche 944</td>
<td>Saab 99</td>
<td>Buick Riviera</td>
<td>Buick LeSabre</td>
</tr>
<tr>
<td>Acura Legend</td>
<td>VW Cabriolet</td>
<td>Olds Cutlass Supr.</td>
<td>Nissan 200SX</td>
<td>Toyota Supra</td>
<td>Porsche 928S</td>
<td>Volvo 740 GL</td>
<td>Toyota Supra</td>
<td>Olds 88</td>
</tr>
<tr>
<td>Sterling</td>
<td>Cadillac Allante</td>
<td>Ford Thunderbird</td>
<td>Toyota Celica</td>
<td>Nissan 300ZX</td>
<td></td>
<td>Volvo 760 GL</td>
<td>Jaguar XJ-6</td>
<td>Dodge 600</td>
</tr>
<tr>
<td>Audi 4000</td>
<td>Mercedes 560SL</td>
<td>Mercury Sable</td>
<td></td>
<td>Buick Somerset R.</td>
<td></td>
<td>BMW 600</td>
<td>Mercedes 190E</td>
<td>Pontiac 6000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Olds Calais</td>
<td></td>
<td>Olds Calais</td>
<td></td>
<td>Jaguar XJ-6</td>
<td>Caddy Eldorado</td>
<td>Ford Taurus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Olds Ciera</td>
<td></td>
<td>Olds Ciera</td>
<td></td>
<td>Mercedes 380</td>
<td>Lin. Mark VII</td>
<td>Mercury Sable</td>
</tr>
</tbody>
</table>

*( - - ) Are Z statistics; see text for details.

Z for 5% significance is 1.65,
1% significance is 2.33.
TABLE 2
Consideration Probabilities, By Category

| Category (i) | Total # Cards | # cards eliminated by categorization | P(E)^j | # cards remaining | P(\bar{E})^j | # cards considered | P(C)^j | P(C|\bar{E})^j |
|--------------|---------------|-------------------------------------|--------|------------------|-------------|-------------------|--------|----------------|
| 1            | 2015          | 1061                                | .527   | 954              | .473        | 235               | .117   | .246           |
| 2            | 1539          | 695                                 | .452   | 844              | .548        | 206               | .134   | .244           |
| 3            | 1509          | 937                                 | .621   | 572              | .379        | 128               | .085   | .224           |
| 4            | 1822          | 1024                                | .562   | 798              | .438        | 179               | .098   | .224           |
| 5            | 2616          | 1162                                | .444   | 1454             | .556        | 405               | .155   | .279           |
| 6            | 1460          | 503                                 | .344   | 957              | .656        | 314               | .215   | .328           |
| 7            | 3592          | 1295                                | .360   | 2297             | .640        | 638               | .178   | .278           |
| 8            | 3510          | 1537                                | .438   | 1973             | .562        | 473               | .135   | .240           |
| 9            | 6030          | 3575                                | .593   | 2455             | .407        | 373               | .062   | .152           |

Notes

P(E)^j = probability of elimination of car in category j = # cards eliminated/# total cards.

P(\bar{E})^j = probability of car not being eliminated in category j = 1 - P(E)^j = # cards remaining/total cards.

P(C)^j = probability of consideration of car in category j = # cards j considered/total cards.

P(C|\bar{E})^j = probability of consideration of car in category j given that car is not eliminated = # cards considered/#cards remaining = P(C)^j/P(\bar{E})^j
(i.e. they were eliminated from further consideration by categorization), leaving 954 of these cards in non-eliminated piles. The probability of elimination $P(E)$ is .527 ($1061/2015$) and the probability of non-elimination $P(\bar{E})$ is .473 ($954/2015$). Of these 954 category 1 cars placed in consideration piles, only 235 were themselves considered seriously for purchase and $P(C)$ was .117 ($235/2015$). The probability of consideration for a category 1 automobile given non-elimination $(P(C|\bar{E}))$ was 0.246 ($235/954$).

This method of calculating a conditional consideration probability was repeated for all nine categories. The probability of initial elimination, $P(E)$, ranges from a low of .344 (category 6) to a high of .621 (category 3). Automobiles identified initially as belonging to category 3 are therefore quite likely to be screened out by individuals in our sample without further evaluation, while category 6 automobiles are particularly likely to receive further attention. The values of consideration given that a car is not eliminated, $P(C|\bar{E})$, range from .152 in category 9 to .328 in category 6. This suggests that category 9 cars are less attractive to this sample of consumers, even when they are not eliminated outright, while category 6 cars are relatively more attractive. The overall consideration $P(C) = P(C|\bar{E}) \times P(\bar{E})$ varies from a low of .062 to .215 or by a factor of more than 3.

Clearly sales will be significantly affected by how customers categorize a new car.

Reatta Response

The automobiles with which individuals placed the Reatta after base case concept exposure were examined. For each individual, the Reatta was assigned to one of the nine aggregate categories identified previously using the assignment rule based on an agreement between the individual and aggregate categories (see earlier description and
The number of people perceiving the Reatta to belong to each of the nine aggregate categories is given in column 2 of Table 3, and this is converted into a probability in column 3 of the same table. Most people (43.5%) viewed the Reatta as similar to group 5 cars given concept exposure. The overall probability of consideration \( P(C)^i \) for each category is taken from Table 2, and are then multiplied by the probability that the Reatta is identified with that category \( P(j) \). The resulting probabilities \( P(C)^r \) are the weighted consideration probabilities, by category. These were summed to provide a total probability of consideration for the Reatta of 14.7\%

Next we look at the probability of dealer visit for the Reatta given that the car is considered. Table 3 shows the average probabilities of visit for those individuals who were assigned to each aggregate category and who would consider the Reatta. This value varies from .23 for category 8 to .63 for category 3. This suggests that categorization affects the probability of an individual searching for more information about the new car in addition to influencing the elimination and consideration processes.

The net effect of the categorization, consideration, and visit phenomena is shown in the probability of consideration and visit \( P(C,V)^r \) for the Reatta. The value varies from .001 in category 1 to .034 in category 5. Assuming the final buying decision is conditioned only by dealer visit and word of mouth communication after visit as we do in our model in Figure 2, the consideration and visit probabilities imply that almost one

---

5 This agrees reasonably well with the average of individual self-assessed probabilities of purchase for Reatta after concept exposure (18.3%) and suggests that the aggregation model is acceptable given that probability of purchase is an adequate surrogate for consideration at the concept exposure level of awareness.
### TABLE 3

Determining Reatta Consideration and Dealer Visit

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>0.014</td>
<td>0.117</td>
<td>0.002</td>
<td>0.50</td>
<td>0.001</td>
<td>1.7%</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>0.035</td>
<td>0.134</td>
<td>0.005</td>
<td>0.62</td>
<td>0.003</td>
<td>4.3%</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>0.028</td>
<td>0.085</td>
<td>0.002</td>
<td>0.63</td>
<td>0.001</td>
<td>1.4%</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>0.052</td>
<td>0.098</td>
<td>0.005</td>
<td>0.29</td>
<td>0.001</td>
<td>1.4%</td>
</tr>
<tr>
<td>5</td>
<td>126</td>
<td>0.435</td>
<td>0.155</td>
<td>0.067</td>
<td>0.50</td>
<td>0.034</td>
<td>49.3%</td>
</tr>
<tr>
<td>6</td>
<td>47</td>
<td>0.162</td>
<td>0.215</td>
<td>0.035</td>
<td>0.53</td>
<td>0.019</td>
<td>27.5%</td>
</tr>
<tr>
<td>7</td>
<td>20</td>
<td>0.069</td>
<td>0.178</td>
<td>0.012</td>
<td>0.51</td>
<td>0.006</td>
<td>8.6%</td>
</tr>
<tr>
<td>8</td>
<td>23</td>
<td>0.079</td>
<td>0.135</td>
<td>0.011</td>
<td>0.23</td>
<td>0.002</td>
<td>2.9%</td>
</tr>
<tr>
<td>9</td>
<td>37</td>
<td>0.128</td>
<td>0.062</td>
<td>0.008</td>
<td>0.27</td>
<td>0.002</td>
<td>2.9%</td>
</tr>
<tr>
<td>Totals</td>
<td>290a</td>
<td>1.00</td>
<td>0.147</td>
<td>0.069</td>
<td>1.00</td>
<td>1.00</td>
<td>100 %</td>
</tr>
</tbody>
</table>

**NOTES:**

a The number of observations used in this table is 290 rather than 434. Although there were 434 clinic participants in total, only 290 were directly involved with the Reatta. The remaining individuals examined comparable, already available models in order to provide control measures for the Reatta forecasts.
half of the sales will come from people who categorize the car in group 5 (see the last column in Table 3) and three quarters from groups 5 and 6.

**Model Forecasting and Simulation**

The aggregate new product forecasting model in Figure 2 can be used to forecast the base case sales and simulate the implications of alternate positioning strategies. The base case is reflected in the categorization and visit probabilities in Tables 2 and 3, the remaining purchase probability and word of mouth flow rates, and the marketing plan in terms of advertising spending, dealer incentives, and pricing (see Urban, Hauser, and Roberts (1990) for discussion of other inputs to the flow model and detailed model equations). Table 4 shows this base case sales forecast for four years, totalling 71,500 units.

As noted above, categories 5 and 6 account for most of the sales in the base case positioning ("a new kind of sports car" with comfort and driveability). Most consumers typify category 5 by the Nissan 300ZX and category 6 by the Porsche 944, so positioning the new car in this group of consumers is the most likely path to high sales. On the other hand, positioning the Reatta strongly as a personal luxury car by Buick and not a sports car could lead to categorization in group 8 along with the Buick Riviera or other existing Buicks. This positioning as a smaller personal sized but typical Buick luxury car can be simulated by changing the fraction of customers who categorize the car into each class (P(j)). For this run, the probabilities for each class were based on the number of
## TABLE 4
Simulated Sales Forecasts for Alternate Positioning Strategies*

<table>
<thead>
<tr>
<th>Positioning Strategy</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Year 4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base case - &quot;New Kind of Sports Car&quot;</td>
<td>11.0</td>
<td>16.3</td>
<td>20.3</td>
<td>24.0</td>
<td>71.5</td>
</tr>
<tr>
<td>&quot;Typical Buick&quot;</td>
<td>6.6</td>
<td>8.8</td>
<td>10.3</td>
<td>12.0</td>
<td>37.7</td>
</tr>
<tr>
<td>&quot;Sporty Riviera&quot;</td>
<td>6.1</td>
<td>8.0</td>
<td>9.2</td>
<td>10.7</td>
<td>34.0</td>
</tr>
<tr>
<td>&quot;Performance Sports Car&quot;</td>
<td>12.4</td>
<td>18.8</td>
<td>23.7</td>
<td>27.9</td>
<td>82.9</td>
</tr>
</tbody>
</table>

* Individual numbers are coded by a constant but they can be compared across strategies.
Buicks already assigned to that category (see Table 1 for "typical Buick" row) and sales decreased dramatically from 71,500 to 37,700 units over four years.

Another positioning possibility was to identify the Reatta with the sportier Buick Rivera and was simulated by increasing the category 8 (where the Riviera is placed) probability to .92 and the other 8 category probabilities at .01. The sales were a little lower than the personal luxury Buick simulation with 34,100 units sold over four years. The Reatta and Riviera are built with the same platform (frame and suspension) and power train (engine and transmission) but the different body and interior two seat arrangement of the Reatta lead it to be categorized in the data as very dissimilar from the Riviera. Using advertising to position the Reatta as a two seat Riviera is not likely to be successful based on this study.

The simulations indicate that the categorization as a sporty car could be successful for Reatta. To simulate the upside potential of an even more sporty positioning we changed the probabilities of categorization as follows: \( P(j=1) = P(2) = .025, P(3) = P(8) = P(9) = 0, P(4) = .075, P(5) = .425, P(6) = .350, \) and \( P(7) = .100. \) These probabilities increase categorization with Porsche, BMW, and Mercedes cars. Sales increased to 82,900 units over the four years. These simulations suggest the Reatta be positioned as a performance sports car from Buick to achieve the highest sales potential, if the car can fulfill the customer's performance expectations.

---

These probabilities were determined by examining the placement of the Reatta car by 40 clinic participants who were shown an initial concept board for the car that identified it as extremely sporty rather than the base case new kind of sports car concept board.
Managerial Implications

Breaking apart the elimination and consideration steps of the decision process suggests that managers have two problems. First, the car must be categorized so it is not eliminated, and second it must have attributes that lead to it being considered given that it is not eliminated by categorization. In the Reatta case, 45.5% of the clinic participants eliminated the car by categorization (calculated by multiplying $P(E_j)$ from Table 2 and $P(j)$ from table 3 and summing over $j$), and two thirds of those who did not eliminate the car initially would not consider it. These are significant numbers, particularly when you recall that the sample was screened to include only people who would consider a two seat car and would consider spending over $20,000. Presumably the elimination was not because of size or price. Positioning communications to help individuals classify the car in either category 6 (Porsche 944) or 7 (where the most typical car is the BMW 300 series) will lead to the least elimination.

Further qualitative research would be valuable to find out why customers eliminated the Reatta from consideration. Consideration given no elimination is highest in categories 5, 6, and 7, but the categorization probability is lower for category 7, with a resulting lower overall consideration probability ($P(C^j)$).

Perceptual maps are one tool for further diagnosing these probabilities. In our case we measured the consideration set on 26 attributes and factor analyzed to find three dimensions which are: 1. "sporty", 2. "luxury", and 3. "reliability" (the heavy loading attributes for the factors were: 1. fun to drive, sporty, and exciting, 2. luxurious, safe, and comfortable, and 3. reliable, good value for the money, and economical, with Eigen
values of 12.2, 3.7, and 1.8 respectively). We then developed perceptual maps for each category. In the target classes 5 and 6 the Reatta was viewed as a little less sporty, more luxurious, and less reliable than the typical class 5 car (300ZX) and class 6 (Porsche 944) so an ad campaign that positions the Reatta as a sporty car and emphasizes its quality and convertible styling could improve the results in the base case. Given the large number of individuals identifying the Reatta with these categories, this strategy would be worth particular attention.

Positioning for categorization, consideration and choice will not be effective if a product does not fulfill its claims. In this study we obtained categorization and rating data after drive of the prototype car as well as after base case concept exposure. The percent of people who placed the Reatta's picture card in each pile after drive was not significantly different than the concept data, but the overall ratings on attributes increased significantly (at the 10% level) on acceleration, handling, comfort, and roomy interior. Word of mouth communication by buyers can be expected to be positive because the car fulfills or exceeds its expectations. Quality in the car production was based on the "craft center" concept where a group of workers assembled the car at work stations rather than on a production line and each worker personally signed the work order to attest to its adherence to specifications. Managers felt this produced a very high quality car. This suggests a more aggressive set of claims in the initial ads would be supported by the actual product quality.

While the overall response to drive performance was good, respondents who categorized the Reatta in class 6 (Porsche 944) and 7 (BMW 300) rated it lower on
performance than the typical cars in these classes. This suggests that attempts to
position the Reatta as a high performance car are not likely to succeed. The model
simulation in Table 4 for "performance sports car" is not likely to be feasible based on
the responses to driving the Reatta.

In conclusion, the categorization, elimination, consideration, and choice modeling
suggested that the car be positioned as a new kind of sports car that offers comfort and
easy handling rather than as a personal luxury car, that Buick's name be secondary to the
sporty image, and that the ads emphasize the car's quality. With this marketing strategy
the model predicts sales of over 70,000 units in four years.

FURTHER RESEARCH AND CONCLUSION

The sub model and measurement methodology for understanding the
categorization, elimination, and consideration consumer decision steps has been
described here in the context of a specific durable goods new product model, but it can
be used or adapted whenever a model is conditioned on evoking or specific brand
awareness. For example, the submodel could be easily added as a front end to new
product trial and repeat purchase and Logit choice models in the packaged goods area
(e.g. SPRINTER by Urban 1970, TRACKER by Blatberg and Golanty 1978, ASSESSOR
by Silk and Urban 1978, and NEWS by Pringle et. al. 1982). Research could be done on
merging the categorization approach with Bass type life cycle models (Mahajan, Muller,
and Bass 1990). An initial approach in this case would be to estimate the three
parameter life cycle model for each category based on customer sorting inputs.
Our model shows sales will vary substantially based on consumer categorization of a new product, so another area of research interest is the identification of categorization cues and their influence on the categorization process. Individual consumers might be asked to verbally describe in detail the typical car identified for each pile formed and why specific piles were formed and eliminated. In addition one could make changes in attribute information for the new product in a more systematic fashion than is employed here in order to assess the impact that individual pieces of attribute information may have on categorization.

A major issue raised in this work revolves around the quality of the aggregation represented in the model. While individual differences in categorization, elimination, and consideration are explicitly incorporated into the proposed model, the discussion here assumes that there is only one true market structure. For example, in the case of the Reatta application, all individuals are represented as perceiving 9 distinct categories of automobiles. Alternatively a full individual model of the process could be proposed and then individual choices added up at the end. Such a micro analytic simulation could be built and compared to the aggregate process model proposed here. Work along these lines has been initiated at MIT. This new methodology will expose potential consumers to a detailed set of information about both existing and proposed new products, and will keep track of the pattern of information examined by each individual, using a multimedia computer system (PC and video disk). Individuals will go through essentially the entire automobile purchase process, but at a much more rapid rate than they would in real life. This new methodology is termed "information acceleration". Measures similar to those
reported here will be obtained at the individual level, and all of this information will then be used as inputs to a micro-level simulation which can be used to investigate individual as well as aggregate consumer responses to a wide variety of different product and positioning options.

The Buick Reatta coupe was introduced in the 1989 model year but production problems restricted its sales to low levels. In the spring of 1990 the Reatta was relaunched with a convertible model and a new advertising campaign based on quality. This campaign showed individuals in the craft center assuring quality in every step of assembly and the sleek styling of a black Reatta convertible. It will be worthwhile to compare the sales forecasted in Table 3 with the actual results over the next four years (see Urban, Hauser, and Roberts 1990 for a description of validation of the original model for another Buick car).

The model described in this paper represents an attempt at integrating behavioral science with quantitative managerial modeling. The use of behavioral science theories in management science models undoubtedly presents researchers with many problems that they might not otherwise face, but the potential promise of such integrated efforts is great. Much more work needs to be done in bringing together the behavioral and quantitative models to improve marketing decision making.
The following is a hypothetical example of how one individual forms several piles, with a discussion of how the information collected from these piles can be used to reveal that individual's categorization, elimination, and consideration processes. In this case, the individual was handed a deck of twenty cards, with the choice alternatives numbered from one to twenty. The familiarity set for this individual consisted of seventeen alternatives (numbers 1 through 17) and the cards representing these alternatives were then sorted by the individual into five piles (see Figure A.1). The resulting piles varied in size from a minimum of two alternatives to a maximum of five. For this individual, the consideration set consisted of the following five alternatives: 1, 3, 4, 7, and 14. Note that this consideration set spans four of the five piles, and that one of the piles contains two consideration set members while three other piles each contain a single consideration set member. If a "consideration pile" is defined to be any pile containing one or more members of the consideration set, then four of the five piles formed by this individual are consideration piles, while pile 5 is eliminated. Following the logic of the decision model described in the text, a new product categorized as belonging to any one of the four consideration piles would be examined further, while a new product categorized in the fifth pile would be eliminated from further consideration.

For the individual's responses represented in Figure A.1, equations (1) through (4) in the text result in the following calculations:
*SHADED BOXES REPRESENT CARS IN THE INDIVIDUAL'S CONSIDERATION SET.
\[ P(E) = \frac{3}{17} = 0.176, \]
\[ P(\bar{E}) = 0.824, \]
\[ P(C|\bar{E}) = \frac{5}{17-3} = 0.357, \]
\[ P(C) = P(C|\bar{E}) P(\bar{E}) = 0.294. \]

That is, an average new brand in this hypothetical product category has an eighteen percent chance of being eliminated outright, a thirty-six percent chance of consideration given non-elimination at the outset, and therefore an overall consideration probability of twenty-nine percent.

This assumes, however, that all categories are equally likely to be chosen for the new product. It is more likely that a consumer will perceive some categories to be more appropriate for the new brand than others. That is, it is necessary to keep separate track of each of the categories formed by the individual. In this case, the category indicator \( j \) would range in value from 1 to 5. For the information represented in Figure A.1, equation (5) in the text would then yield the following consideration probability:

\[
P(C) = \frac{5}{5}(\frac{2}{5})(\frac{5}{17}) + (\frac{3}{3})(\frac{1}{3})(\frac{3}{17}) + (\frac{4}{4})(\frac{1}{4})(\frac{4}{17}) + \\
(\frac{2}{2})(\frac{1}{2})(\frac{2}{17}) + (\frac{0}{3})(\frac{0}{3})(\frac{3}{17}) \\
= \frac{2}{17} + \frac{1}{17} + \frac{1}{17} + \frac{1}{7} + 0/17 \\
= 0.294
\]

This is the same probability of consideration calculated above. Note, however, that it is now possible to see that cars placed in pile 1 are more likely to receive consideration from this individual than those placed in any of the other piles. While this
information appears obvious at the individual level, it is far from obvious once aggregation across individuals occurs.

Now consider the aggregate cluster results depicted in Figure A.2, based on the same product class example used in Figure A.1. Note that while the individual represented in the first example was familiar with only seventeen of the twenty available products, the aggregate clusters include all twenty alternatives. It is also worth emphasizing both that the number of aggregate clusters in this case is smaller than the number of piles formed by the individual as shown in Figure 3 (although for other individuals the number of piles formed could be larger), and that the composition of the aggregate clusters can be considerably different than the composition of any single individual's piles.

Using the aggregate cluster (or category) information in Figure A.2, the assignment rule described in the text can be employed for the person represented in Figure A.1. It should be clear by now that this assignment will be dependent upon which of the five piles the individual places the card for the new product. For example, if the card is placed in pile 1, it is perceived as similar to two category 1 alternatives (item numbers 3 and 4), two category 2 members (numbers 10 and 11), and one category 4 member (number 16). Using the assignment rule, the new alternative would be randomly assigned for this person to either aggregate category 1 or aggregate category 2 (with the probability of selecting each category set at $P = 1/2$). On the other hand, if the card for the new product is placed in pile 2, it is perceived as similar to one category 1 item (number 2) and two category 2 items (numbers 6 and 7). The new product and
## Figure A.2

**Hypothetical Aggregate Categories**

<table>
<thead>
<tr>
<th>Category 1</th>
<th>Category 2</th>
<th>Category 3</th>
<th>Category 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
this individual would in this case be assigned unambiguously to aggregate category 2. Similarly, if the individual places the card for the new product in piles 3, 4, or 5, the new item will be assigned (for that individual) to aggregate category 2, to either aggregate category 2 or aggregate category 3 (.5 probability for each), or to one of aggregate clusters 2, 3, or 4 (P = 1/3 for each) respectively.
REFERENCES


