Information processing models of consumer behavior

James R. Bettman

JMR, Journal of Marketing Research (pre-1986); Aug 1970; 7, 000003; ABI/INFORM Global
pg. 370

INTRODUCTION

So that better decisions can be made by marketers within the context of marketing systems, decision making by firms and by consumers must be understood. Descriptive models of consumer behavior, one aspect of this problem, are the focus in this paper. These are information processing models of individual consumers' grocery product shopping decisions. Information processing models have been successfully applied to other areas of decision making in economics [4, 8], although such models of the behavior of particular individual consumers have received little research effort (see [1] for a model of shopping for women's clothing, and [9] for a model framework, however).

Following Newell, Shaw, and Simon [10], the models considered in this paper have: (1) a memory consisting of an array of cues; (2) a number of simple processes that operate on the cues and develop mediating constructs; and (3) a network, or discrimination net, which represents rules for combining the cues. A decision process is thus viewed as a net through which an array of cues passes. Alternatives are taken at the choice points in the discrimination net depending upon the

*James R. Bettman is Assistant Professor, Graduate School of Business Administration, University of California at Los Angeles. He is grateful to the Ford Foundation and the Richard D. Irwin Foundation for financial support. This study was part of Doctoral research done at Yale University, Department of Administrative Sciences. Funds were also provided by the Division of Research of the Graduate School of Business Administration at UCLA.
Figure 1
THE MODEL FOR CONSUMER C1

KEY TO FIGURE 1

Dictionary: A: Accept
R: Reject
AR: Associate risk (bad experience) with this product
Y: Yes
N: No.

X1: Is this meat or produce?
X2: Is price below justified level?
X3: Is color okay?
X4: Is this the biggest "okay" one?
X5: Is this eggs?
X6: Is the price of extra large over 5 cents more than the price of large?
X7: Is this large size?
X8: Is this extra large size?
X9: Was this product bought last time for this product type?
X10: Was experience with it okay?
X11: Is risk associated with this product (bad experience)?
X12: Is this product class high risk?
X13: Do children or husband have a specific preference?
X14: Is this their preference?
X15: Is it the cheapest size?
X16: Does this class have health (hygiene, diet) factors?
X17: Is this okay on these factors?
X18: Is this for company?
X19: Is the cheapest brand good enough?
X20: Is this the cheapest?
X21: Had a good experience with any brands in this class?
X22: Is this that brand?
X23: Is this the cheapest national brand?
X24: Are children the main users?
X25: Did they state a preference this week?
X26: Have they used this up in the last two weeks?
X27: Is this cheapest size?
X28: Is this that one?
X29: Is this the cheapest size?
X30: Are several "okay" brands cheapest (that they have in stock)?
X31: Is this the cheapest (that they have in stock)?
X32: Have a coupon for this one?
X33: Is this one biggest?
X34: Is there a single national brand?
X35: Is this it?
X36: Have I used this before?
X37: Is this the closest?
X38: Does this feel okay?
X39: Is this for a specific use?
X40: Is this size okay for that?
X41: Is this produce?

value of the cue which that choice point processes. Finally, cues fall into three basic categories: (1) choice object attributes: for example, color, price, weight; (2) external environmental attributes: in order to limit the scope of the modeling process, such complex matters as husband or child preference, use experience with a product, or word of mouth are taken as cues processed by the housewife, but are not explained by a detailed model of their own; (3) internal cues or cognitive variables: the major cue measured here is the degree of risk felt toward a product class [2].

Given this viewpoint, if a model of an individual's processes is desired, how one infers the structure of these processes is an important question. Also, this viewpoint means that one cannot consider many individuals, but rather a few in detail. Finally, procedures for measuring the cues processed must be specified, as in the following section.

RESEARCH METHOD

The researcher followed five housewives (C1–C5) with a portable tape recorder as they shopped for grocery products, each over a six to eight-week period. The researcher knew and selected nonrandomly all the housewives, who were encouraged to think aloud as they shopped. For each choice made, the brand, size and number of packages, price, and aisle position were recorded. The data for the choice object and external environment cues come from these records and taped protocols. The only internal cue measured was the degree of risk a housewife perceived in buying a product...
Figure 2

SIZE SELECTION AND LEGITIMIZATION MODELS
FOR CONSUMER C₁

KEY TO FIGURE 2

Size selection
N: No
Y: Yes
SP: Size positive
SN: Size negative

X35: Is size a storage problem?
X36: Is this for a recipe or specific use?
X38: Will this size handle it?
X39: Is it the right size for this use?

Legitimization
N: No
Y: Yes
LP: Legitimization decision positive
LN: Legitimization decision negative

X40: Has husband mentioned preference?
X41: Hear about it from friends?
X42: Tried some before (at friends or sample)?
X43: Like other products from same brand?
X45: Seen ads on television?
X46: Familiar company source of ads?
X47: Had bad experience with other products from this brand?
X48: Used it satisfactorily previously?

Type. The measurement was ordinal, a rank ordering of product types.¹

Two consumers, C₁ and C₄, were modeled.² Five separate shopping trips were observed for each of these two. The models process the cue array for a given brand and accept or reject that particular brand, given that a product of that type is to be purchased. The models are limited and do not attempt to "menu plan," but take as input that certain product types are to be purchased. Also, the models are restricted by the definition of the external cues. Finally, some of the simple list processing is done external to the model—for example, maintaining the list of products in the store at any given time. These restrictions should give insight into processing of cues and make modeling feasible.

CONSUMER BEHAVIOR MODELS

Consumer C₁

Consumer C₁ had training in mathematics. She has five children. Her husband recently finished medical

¹ A paired comparison questionnaire using 16 product types was used. For every pair of product types A and B, the subject was asked to judge "For which of these two product types would you rather buy a brand you knew nothing about, if you had to buy one of them?" Additional instructions were given to try to ensure a pure choice of the lower perceived risk type—i.e., not a choice of which type is preferred. Finally the 16 product types were ranked according to their scores, with a product type chosen many times a low risk product, and one chosen a few times a higher risk product. For more details, see [3, pp. 36–50]. Also, for a different method of measuring perceived risk, see [7].

² Models were not constructed for C₃ and C₄ because they moved; C₂ was not modeled because she seemed to have the most trouble articulating while she shopped.
X1: Is this product?
X2: Is the color okay?
X3: Is this for a recipe?
X4: Is it small enough?
X5: Is this the right size?
X6: Does it feel okay?
X7: Had experience with this product type?
X8: Is there a convenience type feature?
X9: Does this have that feature?
X10: Can you see the product itself?
X11: Does this have the best color?
X12: Are packages (cans or boxes) damaged?
X13: Does this have least damage?
X14: Is size satisfactory?
X15: Is this the cheapest "okay" one for given size?
X16: Is this high perceived risk?
X17: Had bad experience with this brand?
X18: Is this one legitimized?
X19: Is it the only brand available?
X20: Is it really necessary to have it now?
X21: Does husband have strong preference for brand?
X22: Is this that brand?
X23: Does this product type have a convenience feature?
X24: Does this one have it?
X25: Is there a freshness factor?
X26: Is this factor necessary for this end use?
X27: Is this brand sufficient for this end use?
X28: Does this have that factor?
X29: Was this brand bought last time?
X30: Was last brand liked better than others previously tried?
X31: Are they out of that brand?
X32: Did legitimization of this brand involve trial?
X33: Was this one better than the old one in that trial?
X34: Was this liked better than others previously tried?
X35: Any new brands legitimized since then?
X36: Did this involve trial?
X37: Was the old one better in this trial?

The figure above is a diagram representing the model for consumer C1. The nodes and edges in the diagram illustrate the decision-making process for selecting a product. Each node represents a decision point or a piece of information that influences the consumer's choice. The edges indicate the flow of decisions and the influence of one decision on another.

The flow chart of consumer C1's model is in Figure 1. The first branches represent the decision rules for meat, produce, and eggs. For meat and produce, C1 used the highest justified price (X2) as a learned trigger level for choosing a cut of meat or type of produce, and then processed the primary sensory cues of color, heaviness, and feel (X3, X4, X43, X44). For eggs (X5), the rule shown is one C2 said she remembered reading in Consumer Reports (although the parameter she used, five cents, may not have been accurately remembered).

The nodes X9, X10, and X11 represent C1's feedback from experience. If it was bad (X10), she would associate risk with that brand and reject it later (X11). (This is not the same as perceived risk, but represents a rejection flag on a particular brand.)

A major branching point in the model is node X12, the perceived risk cue. For a high risk product class (high on the perceived risk ordering), C1 checked if her husband or children had a preference (X13), or if the class had generalized health factors (X16; e.g., hexachlorophene soap or sugar-free canned fruit). If so, she bought the cheapest brand which satisfied the preference or health factor (X14, X15, and X17, X30, X31).

For a low risk product, C1 departed from buying the cheapest only in certain cases, for guests (X18, X19) or the children (X24, X25, X26, X27, X28, X29). She still tried to buy the cheapest within these constraints (X23 and X27 or X29). Otherwise, with no such departures, she bought the cheapest of the "okay" products in stock at the time, in terms of price per unit. This list of products with no associated bad experience (as in X10 or X17) is maintained external to this model.

Since the model for C1 is largely price-oriented, how she codes products with respect to certain cues (e.g., X18, X24, X13, X16) will determine the number of products she considers before making a decision.
exceptions to the rule of buying the cheapest. The protocol states:

I get the, the little tenderized ones [steaks], what are they called? These are the ones, if they get below a dollar (X2).

By what color it is . . . Every picture of watermelon one sees has a picture of pink, quite pink watermelon (X3).

I picked it because it was the biggest one [pork shoulder] (X4).

Well, it depends on the price differential. If there's less than what I feel, I can't buy them, I've just not nutted, when I get the ones that are supposed to be extra savings (X10, X11).

The children, I don't, or my husband. They said it was better (X13).

My daughter . . . feels that a deodorant soap or hexachlorophene or whatever it is is good for her face . . . so I get the cheapest deodorant soap (X16, X17, X30).

This part I really can't do without a child to tell me which cereal is in favor this week (X25).

I came last week and divided everything by 14 ounces into 59 and 12 ounces into 49 and so forth and finally figured out which is the best. Now until they change the price on me, I'm all set . . .

It's a little exhausting by the time you get through (X31).

Consumer C4

Consumer C4 presents a vivid contrast to consumer C3. She is younger and newly married, with no children. She seems to be more risk in making product decisions than C3. She likes to stick with what others had tried and told her to try. Her husband has strong preferences (he didn't like to try something he had not had before) and thus influence was bought. Other factors in her decision process were desires for convenience and freshness:

My theory, whatever works . . . If it's something that I like, for two or three times, then I'll keep on getting it rather than change, usually . . . I'd still rather [get] this, because I've used it before [flour]. . . Most of them have only tomatoes, but they have celery and everything in it so you don't have to put anything else in them. . . We like everything very fresh, and we'd rather go and get some more at another time later in the week than buy a larger quantity now . . .

The flow chart of the model for C4 is in Figures 2 and 3. Figure 2 shows the processes of legitimization and size selection. The outputs of these are then used as mediating cues in the main process shown in Figure 3 (i.e., X14 for size selection and X18 for legitimization are set to Yes or to No in the main model depending upon the outputs in Figure 2).

If in a size decision there is no storage problem (X35) or no recipe or specific use (X36), C4 bought the smallest size for freshness (X37).

Legitimization is the process by which a product becomes eligible for C4's consideration. Since she usually did not try brands she did not know, the major cue types are her past experience (X48); word of mouth (X40, X41, X42); trial (X43); or marketing influences (X44, X45, X46). These inputs (except X48) are conditioned, however, by any negative halo effect (X47).

Now consider the main model in Figure 3. Since C4 bought meat elsewhere, the first branch considers only produce. Here color and feel must be satisfactory (X2 and X6); size is determined by end use (X3, X5) or freshness considerations (X4).

If C4 has not had any experience that she can recall with a product type (X7), she tries to reduce the perceptual field by using convenience as a filter property (X8, X9): then, if the product itself is visible, she judged by color and size (X10, X11, X14); if it was not visible, she chose by condition of the containers (X12, X13) or price (X15), if size was satisfactory (X14).

If she has had experience with this product class, the next major branching occurs for the perceived risk cue (X16). If a product was low risk and had no stigma of bad experience (X17), then the cheapest one of satisfactory size was selected (X15). However, if a product was high in perceived risk and not legitimized (X18), it was bought only if necessary at that particular time (X20) and the only brand available (X19).

Otherwise, if the product was legitimized, C4 considered her husband's strong preferences (X21). She next looked for a convenience feature, if necessary (X23, X26, X27), or a freshness factor (X25, X28). If the product passed all these screening cues, the main process of the model is entered, the habit or loyalty process. If the brand was purchased the previous time and is still preferred to others tried before (X29, X34), and no new brands have been legitimized and preferred in a trial (X49, X50, X51), then the brand bought last time is bought again, by size (X14). If the brand was not purchased last time, but the last brand was disliked (X30); the retailer is out of the last brand bought (X31); or if this new brand has been legitimized by trial and is better than the former brand in that trial (X32, X33), then the new brand is purchased by size (X14). Consumer C4's comments illustrate the processes:

I get the small kind, because it's got a wide mouth, so you can get all of it out of it, and I have a very small refrigerator, and the bigger bottles don't fit in (X35, X23).

If it's not friends or on television, forget it (X42, X45). We tried some. We were given it free in the supermarket (X43, X32, X50). This one I'll take because it's solider than the other one, and it's also a good dark green (X1, X2, X6).

First I was looking for Bartlett pears, but they both said Bartlett pears, so then I looked to see, um, if one of them, if all the cans were in good condition in one . . . and if one of them, all of them were in good condition, then I'd get that one, over the ones that most of
them were bashed in. The third thing I looked for was price, and they were the same price (X7, X12, X13). I'll get these. By price. We use them up fast, so I might as well get the cheapest thing... paper products don’t matter to me (X16, X17, X15).

I’m getting this kind because this is the kind [husband] said he likes (X21, X40).

This time I’ll get Brand A [not usual one], because they have everything in it,... but they don’t have any Brand B [usual one] with everything in it (X23, X31).

I like this one. We've used it before, and we've had good luck with it (X29, X34).

Consumers C1 and C4 have markedly different model behavior styles. This may be partially due to the differing situational and demographic factors involved and to differing degrees of shopping experience. Cox, with his two subjects June and Marsha, found much the same type of contrast in one of his studies [5].

MODEL TESTING

To qualify as an adequate representation of a process, a model must pass the test of output validation. Although not sufficient, it is necessary to test whether a model can replicate reasonably well the decisions actually made. One problem with output validation for a complex process model is that testing all the decision rules imbedded in the model (even seldom-used rules), requires large amounts of data. However, the data collected previously for C1 involved 226 decisions, but for C4 only 70, and not all of these could be coded, as discussed further below. Also, some cues specified in the model cannot be ascertained—i.e., asking the housewife whether her husband or child prefers Product A or whether color was satisfactory.

Despite these difficulties, it was felt that some attempt at output validation had to be made as one way of gaining confidence in the models. As a result, the following methods are a compromise.

Table 1

PERCEIVED RISK RANKINGS FOR THE CONSUMER MODELS

<table>
<thead>
<tr>
<th>Consumer C1</th>
<th>Consumer C4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Flour</td>
<td>1 Hot dogs</td>
</tr>
<tr>
<td>2 Lettuce</td>
<td>2 Ketchup</td>
</tr>
<tr>
<td>3 Instant coffee</td>
<td>3 Frozen vegetable</td>
</tr>
<tr>
<td>4 Tomato juice</td>
<td>4 Soft drink (six pack)</td>
</tr>
<tr>
<td>5 Canned fruit</td>
<td>5 Canned vegetable</td>
</tr>
<tr>
<td>6 Frozen orange juice</td>
<td>6 Instant coffee</td>
</tr>
<tr>
<td>7 Frozen vegetable</td>
<td>7 Frozen orange juice</td>
</tr>
<tr>
<td>8 Hot dogs</td>
<td>8 Oleomargarine</td>
</tr>
<tr>
<td>9 Canned vegetable</td>
<td>9 Canned fruit</td>
</tr>
<tr>
<td>10 Oleomargarine</td>
<td>10 Tomato juice</td>
</tr>
<tr>
<td>11 Soft drink (six pack)</td>
<td>11 Mayonnaise</td>
</tr>
<tr>
<td>12 Ketchup</td>
<td>12 Flour</td>
</tr>
<tr>
<td>13 Paper towels</td>
<td>13 Eggs</td>
</tr>
<tr>
<td>14 Mayonnaise</td>
<td>14 Lettuce</td>
</tr>
<tr>
<td>15 Eggs</td>
<td>15 Laundry bleach</td>
</tr>
<tr>
<td>16 Laundry bleach</td>
<td>16 Paper towels</td>
</tr>
</tbody>
</table>

Table 2

RESULTS OF THE CONSUMER MODELS

<table>
<thead>
<tr>
<th>Consumer</th>
<th>Original data actual accept/reject</th>
<th>Validation data actual accept/reject</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 Model</td>
<td>Accept 68</td>
<td>Reject 13</td>
</tr>
<tr>
<td>C1 Model</td>
<td>Accept 68</td>
<td>Reject 13</td>
</tr>
<tr>
<td>C4 Model</td>
<td>Accept 41</td>
<td>Reject 5</td>
</tr>
</tbody>
</table>

The models were programmed in FORTRAN with two sets of data: the original shopping data from which the models were derived and a second set taken especially for validation purposes. For the original data, only those past decisions whose cues could be coded from the tapes were included. Of the 226 original decisions for C1, 87 were coded; of the 70 for C4, 50 were coded. This sample (unless specific cases of a product not being purchased were discussed in the protocol) could only test valid positive or false negative judgments.

The second set of choice data, 78 choices for C1 and 46 for C4, was collected about four months after the first set: the consumers shopped, a list of the products purchased was made, and the brand, price, and weight of each were recorded. The consumers were then asked questions from the model until enough cues were specified for the model to be able to make a decision. In addition, for most of the products purchased, unchosen alternatives were recorded at the store (to ensure that they were actually in stock), and the procedure above was followed with the subject until the model could make a decision on each of these alternatives. Finally, high and low risk rankings were determined for those product types not ranked in the original 16-product risk ranking.

Tables 1 and 2 present the findings of the study. In Table 1 the perceived risk rankings are given, with one the highest risk, and 16 lowest. Because C4 did not seem to perceive extreme product risk, the median was used as her dividing line between high and low—Product 8 was high and Product 9 low risk. On the other hand, C4 seemed to feel very great risk in the shopping situation, so her cutoff point was set at the three-quarters point—Product 12 is high and Product 13 low risk. This cutoff point was verified in the tapes, since C4 showed that flour was risky and that paper products do not really matter.

Table 2 shows the results of running the models on the data and how well each model matched the actual decisions for each data sample. The diagonal entries represent correct predictions. The percentages of correct predictions for C1 are thus (a) original data—85.1% (p = .00, 6.54α); (b) validation sample—89.7% (p = .00, 7.02α); and (c) combined samples—87.2% (p =
.00, 9.57σ). For C4 the percentages are (a) original data —88.0% (p = .00, 5.37σ); (b) validation data—87.0% (p = .00, 5.01σ); and (c) combined samples—87.5% (p = .00, 7.35σ). The probabilities listed after each result are those of having that percentage or higher correct, using a normal approximation to the binomial with p = .5. The results are significantly better than the expectations under a null hypothesis of guessing accept or reject with equal probability.

Thus the models predict rather well. For behavioral models, process validation is always an important concern. The excerpts from the protocols can be considered a weak Turing's test approach [4, pp. 77–90]. However, the very process of modeling from tape recorded conversations should ensure at least reasonable performance on Turing's test (see [3, pp. 111–224] for new techniques and theoretical approaches).

**IMPLICATIONS OF THE MODELS**

The models for these consumers match their actual decisions with a high level of precision. Three factors contribute to this predictive accuracy: (1) the models do not "menu plan," and hence their task is easier than it might be; (2) the external cues are very general and afford a lower level of explanation than might be wished; and (3) the process of determining cue values, although attempts were made to keep it as objective as possible, still has many subjective elements. Despite these possible objections, the ability of the models to predict accept-reject decisions was very satisfactory.

The major finding of this study is that the product accept-reject decisions of an individual consumer are modeled quite well by a decision net of cues. Such a net has several major subprocesses. The process of cue perception is very important for determining which of these subprocess decision rules are used dominantly. One shortcoming of the present models is that such factors as attitudes and social elements are presumably important in cue encoding, but have not been explicitly dealt with, although some cues deal with these factors implicitly. The models do not explain how product types, for example, are seen as having a convenience or health factor. The models do, however, focus attention on the cue combination process—and the fact that decisions in the consumer models depend upon configurations of cues justifies this focus.

In addition to these broad findings about the structure of the decision process, the models also imply that perceived risk is an important product cue. The difference in risk ratings by C1 and C4 implies that a marketer could attempt to assess how his product type's risk ranking varied across the target population for the product. Then he can decide, on the basis of this population distribution, how important it is for his brand to attempt to reduce felt risk.

In conclusion, it is hoped that these findings show that decision process modeling by gathering protocols can lead to better understanding of marketing phenomena. Information processing models of aspects of consumer behavior offer the researcher a very rich field of study. Such models can form the basic data points from which a cognitive theory of consumer behavior can be induced.

**REFERENCES**

2. Raymond Bauer, "Consumer Behavior As Risk Taking," in [6], 23–33.