THE COST OF THINKING:
ITS IMPLICATIONS

Steven M. Shugan

INTRODUCTION
Research involving choice from some finite set of alternatives and the associated information processing continues to grow (Bettman, 1978; Einhorn, 1971; Green and Wind, 1973; Hauser, 1978; Keeney, 1974; Simon, 1957). Some psychologists have suggested that the utility maximization may be beyond the limitations of the human mind and suggest that consumers adopt simplifying decision rules. For example, the simplifying conjunctive rule states that any alternative not meeting a minimum cutoff level on any characteristic is immediately eliminated from consideration. The Federal Drug Administration uses a conjunctive rule when issuing standards, such as purity, weight, age, etc., that all ethical drugs must meet. However, it is clear that these rules will often reduce expected benefits by allowing selection of a less than optimal alternative.
The rules essentially ignore some possibly useful information. It is therefore necessary to understand exactly what cost savings, if any, are associated with these rules, so that these savings can be compared against the resulting lower expected benefits.

A recent paper titled, "The Cost of Thinking" (Shugan, 1980) deals with the problem of explicitly considering the cost side of decision-making rather than merely the benefit side. The paper develops a theory by which the potential difficulty of a decision can be measured. This paper explores three specific implications of "The Cost of Thinking." First, the reaction of a decision maker who faces a recurring choice where the previously chosen alternative is no longer available is considered. Second, the paper considers the reaction of a decision maker who faces a recurring choice where a new alternative becomes available. Third, the paper considers effects of specific information supplying mechanisms, i.e., persuasion. The paper closes with several other interesting choice problems, a summary and a discussion of future research issues.

DEVELOPMENT OF A "THINKING COST"

It will be useful to first consider the cost of a utility maximizing model. A consumer who wishes to choose the best (most preferred) alternative from $M$ different alternatives must somehow eliminate the other $M-1$ alternatives. Hence, that consumer must make $M-1$ alternative comparisons in order to find the most preferred alternative. If there was some fixed cost, denoted $f$, for each alternative comparison, the total cost of the decision could be computed by multiplying that cost by the number of alternative comparisons. The difficulty of the choice would be given by the expression $(M-1)f$. This expression provides a method for computing "thinking costs." However, some alternatives may be harder to compare with each other than other pairs of alternatives.

Therefore, we must proceed to develop some measure of the difficulty associated with comparing two alternatives. One method would be to assume the consumer's preferences are determined (directly or as a cue) by the alternative's characteristics. Therefore, a consumer who wishes to choose between any two alternatives may progress by comparing the two alternatives on their characteristics. For example, a consumer choosing between two Chicago restaurants may first compare them on decor. Second, a comparison on menu variety may take place. Next, the restaurants could be compared on service quality. These comparisons could then continue until all alternative characteristics are exhausted, uniquely defining each choice alternative. Undoubtedly, this process can be lengthy and tiresome.

Now, the choice of an alternative is preceded by a number of char-
acteristic comparisons. Assume that associated with each characteristic comparison is a fixed cost: a unit of comparison effort. It may be then reasonable to assume that the more characteristic comparisons necessary to make a choice, the more difficult the choice. We will measure the thinking cost associated with a choice by positing that \( f \) is monotonically related to the number of characteristic comparisons made. That is, more difficult decisions require more characteristic comparisons.

In order to implement this proposition, we will interpret the problem of selecting characteristics as one of sampling. The consumer can be viewed as sampling alternative pair differences by characteristic. For example, consider the Chicago restaurant alternative comparison. The consumer first compares the restaurants on decor. This characteristic comparison is essentially sampling from the population of alternative differences. The sample chosen has one observation: difference in the characteristic of decor. Again, the second comparison on menu variety can be interpreted as an observation on menu variety difference. Further characteristic comparisons can once more be considered new observations in the sample. At each point in the sampling process, the consumer can infer which restaurant would provide the largest satisfaction. This inference could be done by simply considering the sample mean which represents an estimate of the average difference in characteristic utility between the two restaurants.

The crucial question becomes how many characteristic comparisons must take place in order to make the choice decision. The number of characteristic comparisons determines \( f \) and hence the required effort associated with the choice. Three factors influence \( f \) and hence the difficulty of the choice (Shugan, 1980). These factors are:

1. The true difference in utility between the two alternatives. If \( z_j \) represents the difference between the products on the utility of characteristic \( j \), then the true difference in utility is reflected by the expected value of \( z_j \) probabilistically chosen, which will be denoted \( E(z) \).
2. The confidence level at which the decision must be made, denoted \( \alpha \). This value is the probability of not making a mistake. Hence, an \( \alpha \) near one implies a high level of required confidence and involvement in the decision.
3. The variability in the characteristic utility difference between the two alternatives. This quantity is the variance of \( z_j \) over \( j \) and will be denoted by \( \text{VAR}(z) \).

For example, consider Table 1. Table 1 illustrates characteristic utilities for automatic drip coffee makers taken from Consumer Reports (1980).
Table 1. Characteristic Utilities for Drip Coffee Makers

<table>
<thead>
<tr>
<th>Brand and Model</th>
<th>Coffee Quality</th>
<th>Pouring Ease</th>
<th>Cleaning Ease</th>
<th>Expected Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Braun KF-20</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>4.00</td>
</tr>
<tr>
<td>Norelco HB5140</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>3.00</td>
</tr>
<tr>
<td>Krups 261A</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4.67</td>
</tr>
<tr>
<td>Bunn b-8</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>2.67</td>
</tr>
</tbody>
</table>

We assume that these ratings are given equal weight and hence represent utilities; however, any weighting scheme would be satisfactory. In general, data of this type can be obtained by multiplying importance weights by beliefs in a linear multi-attribute model (e.g., see Lutz and Bettman, 1977), or from part worths in conjoint analysis (e.g., see Green and Rao, 1971) or through some intensity measure (e.g., see Hauser and Shugan, 1980). In this example, we can compute the difference in characteristic utility for any alternative comparison. For example, when comparing Braun KF-20 with Norelco HB5140 we find: \[ z_1 = 5 - 4 = 1, \quad z_2 = 5 - 3 = 2 \] and \[ z_3 = 2 - 2 = 0. \] Further, when comparing Braun KF-20 with Norelco HB5140 we find: \[ E(z) = (1 + 2 + 0)/3 = 1.00 \] and \[ \text{VAR}(z) = (1^2 + 2^2 + 0^2)/3 - (1.00)^2 = .67. \] When comparing Braun KF-20 with Krups 261A we find: \[ E(z) = -.67 \] and \[ \text{VAR}(z) = 2.89. \] When comparing Braun KF-20 with Bunn b-8 we find: \[ E(z) = 1.33 \] and \[ \text{VAR}(z) = 9.56. \]

Sampling theory tells us that if the true difference in average characteristic utility is large, fewer characteristic comparisons need be made to insure the same confidence level (holding other factors constant). Fewer necessary characteristic comparisons imply an easier choice. Hence, we would suspect the Braun-Norelco comparison is made less difficult than the Braun-Krups comparison because \( E(z) \) is smaller in magnitude (.67 rather than 1.00) for the latter comparison.

Sampling theory also tells us that if the variability is large, then more characteristic utility comparisons are necessary to insure the same level of confidence (holding other factors constant). Hence, the Braun-Norelco comparison is made even less difficult than the Braun-Krups comparison because \( \text{VAR}(z) \) is far greater for the latter comparison.

Finally, sampling theory tells us that if we require more confidence, we must make more characteristic comparisons. A larger sample of characteristic comparisons, in turn, makes the choice more difficult. If these additional comparisons are not made, the consumer would face an unacceptably high risk (as defined by the confidence level). Remember, the consumer accepts some possibility of making a mistake (i.e., \( 1 - \alpha \)) because all alternative characteristics are not compared.

Before proceeding it is important to mention two points discussed
elsewhere in more detail (Shugan, 1980). First, it is important to realize why the confidence level is determined from outside of the development. The confidence level represents the involvement level of the consumer. It reflects how this choice interacts with all other decisions. Hence, it is partially determined by opportunity costs. As an increase in house maintenance costs incites curtailment of food expenditures, a forthcoming marriage decision may raise the opportunity costs of thinking about the purchase of an automatic drip coffee maker. Further, the relative losses incurred from a mistake may also be greater in the case of a choice of mates than a choice among drip coffee makers. Here, assume $\alpha$ reflects and captures the effect of all outside choices on the choice at hand.

Second, we are, in fact, not measuring the true choice difficulty because that quantity is probabilistic. We are, instead, measuring the expected, or, more accurately, potential difficulty of the choice situation. As researchers or observers, we know the true distribution of characteristic utilities. The consumer does not! We are determining how many characteristic comparisons are necessary to perform the choice task at some confidence level. In essence, we are determining a measure for the relative difficulty of the choice task facing the consumer. Of course, the actual difficulty depends on the luck of the decision-maker.

THE CONFUSION INDEX

Shugan (1980) derives a bound on the required number of characteristics necessary to perform the choice task and calls that bound the confusion index. That bound for a binary comparison is,

$$\frac{\text{VAR}(z)}{(1 - \alpha) [\text{E}(z)]^2}$$

The confusion index provides a measure of decision making difficulty. Further, if thinking effort is held constant across decisions, then $\frac{\text{VAR}(z)}{[\text{E}(z)]^2}$ can be interpreted as a measure of dissimilarity between the two alternatives. Hence, Braun is most similar to Norelco (the easiest comparison) and least similar to Krups (the hardest comparison). We can now proceed to discuss several important theoretical problems where a knowledge of "thinking costs" could prove helpful.

ALTERNATIVE DELETION

The problem of alternative deletion has both practical and theoretical significance. In this situation, the consumer is faced with a recurring choice where the previously chosen alternative is no longer available.
One example of this problem occurs when a gasoline shortage makes one traditional mode of travel infeasible and the consumer must search for an alternative mode of transit. The problem of alternative deletion is also of importance to many researchers because it is often necessary to make some assumption concerning the effect of alternative deletion in order to estimate preference functions from choice data (McFadden, 1970). One convenient way to approach the deletion problem is to postulate some probability distribution over the possible alternatives and some mechanism for updating the distribution in the event one alternative becomes unavailable. For example, Luce's choice axiom (Luce, 1959) assumes all remaining alternatives' choice probabilities gain proportionally from the deleted alternative's choice probability. It is often conceptually easier to think of each choice probability as either the researcher's belief that a particular alternative will be chosen or as an aggregate market probability (i.e., market share) rather than the individual consumer's choice probability. Then, in marketing terms, Luce's axiom assumes proportional gains in market share. For example, if four alternatives exist with market shares of 20%, 8%, 40% and 32%, respectively, Luce's axiom would predict that deletion of the first alternative would result in new market shares of 0%, 10% (i.e., .08/(1 - .2)), 50% and 40%, respectively. However, proportional absorption may be an unreasonable assumption when the choice alternatives exhibit strong differences or similarities. For example, the deletion of a particular French restaurant from the dining choice set may provide more business for other French restaurants than the local truck-stop.

In order to adequately deal with marked differences in similarity among alternatives, it becomes necessary to examine more closely the composition of the choice alternatives. For example, Tversky (1972) defined alternatives as collections of aspects (binary characteristics). In this framework, French restaurants would share the common aspect of specializing in French food while the local truck-stop might lack this aspect. Tversky's mechanism assumes that the consumer sequentially selects aspects and eliminates all alternatives not possessing the selected aspect. The selection probabilities for aspects obey Luce's Axiom. Using Tversky's mechanism requires some knowledge concerning alternative similarities. In the previous four-alternative example, if the deleted alternative's 20% market share came half from aspects unique to the deleted alternative and half from shared aspects, the remaining three alternatives would proportionally absorb the 10 percent share from unique aspects and absorb the other 10 percent share according to the number of shared aspects. However, Tversky's mechanism also has its limitations. For example, the mechanism would allow (perhaps with some low probability) a prudent consumer considering the choice between one of two

Cost_t_implications2.max
houses to select one house over the other house merely because the latter house possesses a leaky faucet. This decision would be made regardless of the price differential (using the price level as one aspect) even though the price differential might easily justify the fixing of the faucet. This weakness comes from the lack of a behavioral rationale for the mechanisms by which attributes are selected. The mechanism also forces the consumer to observe at least as many aspects as there are alternatives, and deals with continuous attributes in an awkward manner.

It is therefore necessary to attack the alternative elimination problem from a more general viewpoint. Consider a consumer faced with the unavailability of the previously chosen alternative. This consumer can proceed with one of four actions:

1. Compare all available alternatives on all their characteristics
2. Compare some available alternatives on some of their characteristics
3. Compare all available alternatives on some of their characteristics, or
4. Compare some available alternatives on some of their characteristics.

If the consumer compares all available alternatives against the deleted alternative or against each other, the choice becomes exceedingly difficult. In this situation, all possible binary characteristic comparisons take place and hence the limit of difficulty is achieved. This action may be elected by the consumer when the required confidence level is extremely high, the number of choice alternatives is relatively small or, perhaps, no other choice decisions are pressing and “thought” is therefore cheap. For example, the situation might dictate the possibility of a great loss associated with a mistake or a great benefit associated with a correct decision. In any case, the study of the exhaustive choice strategy would only become interesting when the full expected benefits of thought are considered and the total opportunity costs associated with competing decisions are analyzed.

If the consumer compares all the available alternatives on some of their characteristics, the process can be viewed as a new choice problem. The consumer must now choose the best alternative from some fixed set of alternatives at some acceptable confidence level where the new choice set excludes the unavailable alternative.

If the consumer compares some of the available alternatives on all of their characteristics, the “confusion index” provides little insight. The problem reduces to a search problem where no individual choice decisions are errored. However, the consumer must now decide which alternatives to consider in the choice process. If the consumer randomly samples from the available alternatives, we would suspect that the size
of that sample would depend on the confidence level at which the decision must be made, the expected variance in alternative utility across all alternatives and the nature of the expected distribution of alternative utility across alternatives. A high required confidence level and a large expected utility variability will, of course, imply a large sample of alternatives should be considered.

Finally, the consumer may compare some of the remaining alternatives on some of the characteristics. Of course, the consumer could again begin the choice problem with no consideration of the now unavailable alternative. However, suppose the consumer uses the previously selected and now unavailable alternative as a benchmark. Then the consumer would proceed by comparing the available alternatives against the previously selected alternative until some available alternative is found with at least the same utility as the previously selected one, i.e., the benchmark. To further generalize the problem, it is also possible to assume that the consumer continues to evaluate choice alternatives until some lesser level of utility is inferred because the consumer views the utility of the unavailable alternative as unattainable or not worth seeking. In this situation, we can compute the difficulty associated with each decision (i.e., the available alternative with the benchmark) by computing \(E(z)\) and \(\text{VAR}(z)\). Here \(z_j\) represents the difference in the utility of characteristic \(j\) computed by comparing the benchmark with the alternative under consideration. Then, the difficulty associated with each alternative comparison would vary depending on the dissimilarity of the alternative with the benchmark alternative. One measure of that dissimilarity, as indicated by the confusion index, would be \(\text{VAR}(z)/(E(z))^2\). This measure provides several interesting implications. For example, consider two of the available alternatives. Suppose the first alternative is superior to the benchmark alternative, that is, \(E(z)\) for the first alternative equals some constant greater than zero. Further, suppose that the second alternative is inferior to the benchmark alternative by the same constant. If the consumer allocates equal effort when comparing each alternative with the benchmark, the consumer is more likely to make a mistake for the more difficult alternative comparison. Further, the alternative comparison involving the larger \(\text{VAR}(z)\) would lead to a greater likelihood of an error. In the decision involving the first alternative, a mistake would mean eliminating a superior alternative from consideration. However, in the decision involving the second product, a mistake would mean choosing an inferior product. Hence, the greater the similarity of the first alternative with the benchmark alternative, the greater the chance the first alternative will be chosen. However, the greater the dissimilarity of the second alternative with the benchmark, the greater the chance the second alternative will be chosen. In general, dissimilar inferior alter-
natives and similar superior alternatives are more likely to be chosen than similar inferior alternatives and dissimilar superior alternatives where dissimilarity could be measured by \( \text{VAR}(z)/(E(z))^2 \).

If we further decompose \( \text{VAR}(z)/(E(z))^2 \) into its components, we find similarity can be equated with,

\[
\frac{\text{VAR}(x) + \text{VAR}(b) - 2 \text{COV}(x,b)}{E(x) - E(b)}
\]

where: \( \text{VAR}(x) \) is the variability in characteristic utility for alternative \( x \).
\( \text{VAR}(b) \) is the variability in characteristic utility for the benchmark.
\( \text{COV}(x,b) \) is the covariance between the characteristic utility for alternative \( x \) and the benchmark, computed across characteristics.
\( E(x) \) is the actual or expected utility (averaged across characteristics) of alternative \( x \).
\( E(b) \) is the expected characteristic utility for the benchmark.

Hence, if the benchmark alternative had high characteristic utility variability, \textit{ceteris paribus}, the likelihood of a mistake is high. In this situation, more similar superior alternatives and more dissimilar inferior alternatives are more likely to be chosen. If the alternative under consideration has high characteristic utility variability, \textit{ceteris paribus}, the likelihood of a mistake is high. However, the larger the covariance between the benchmark and alternative's characteristic utilities, \textit{ceteris paribus}, the lower the likelihood of a mistake. Finally, the larger the difference between the benchmark and alternative average (or expected) characteristic utility, the lower the likelihood of a mistake.

It is now possible to make predictions regarding the distribution of choice probabilities following an alternative deletion. For example, when the benchmark alternative becomes unavailable, the consumer is most likely to choose inferior alternatives with negative or small characteristic covariances, high characteristic variances and small mean characteristic differences. In contrast, the consumer is most likely to choose superior alternatives with large positive covariances, low characteristic variances and large mean characteristic differences. Hence, from a marketing perspective, a competitive but inferior product trying to capture the market share of an absent competitive product would want to convey characteristics it has which dominates the absent product and, at the same time, those characteristics on which the products drastically differ (creating negative covariance). The opposite implications would follow for a competitive but superior product trying to capture the absent product's share.

The previous argument assumes an equal allocation of decision-making
effort across alternatives. However, consumers may allocate effort in proportion to the difficulty of the choice. In that case, the asymmetric relation between superior and inferior alternatives with respect to decision-making error does not exist. The consumer merely allocates more effort to dissimilar alternative comparisons than similar alternative comparisons so as to achieve the same confidence level. And, of course, in the marketing example, the superior product could always adopt an advertising strategy stressing the importance of a full evaluation (e.g., large losses associated with a mistake). In fact, if consumers tend to bias effort so as to overallocate effort to dissimilar alternative comparisons, then the asymmetric error actually reverses in direction. Hence, the confidence level at which individual decisions are made does drastically affect the consequence of alternative deletion.

ALTERNATIVE ADDITION

Another problem facing choice theorists is the problem of alternative addition. For example, marketers interested in the process of new product development and defensive tactics for existing products challenged by a new product frequently encounter the problem of alternative addition. Here, the consumer is faced with the recurring choice among some set of alternatives when some new alternative suddenly becomes available. Again, by using reasoning similar to that of alternative deletion, we find an asymmetric relationship with respect to alternative comparison error when decision-making effort is uniformly allocated. However, in the case of alternative addition, the consumer can generally elect one of the following three actions.

1. The consumer may not evaluate the new alternative.
2. The consumer may evaluate the new alternative with a bias toward rejecting it in favor of the last chosen alternative.
3. The consumer may evaluate the new alternative in an unbiased manner.

In the first case, the new alternative would become simply one of many untried alternatives in the alternative choice set. In order for the new alternative to be considered, some incentive would need to be provided for the consumer to take this action. Whenever the consumer no longer evaluates all possible alternatives in the choice set, it is necessary to consider such factors which initiate evaluation. Some of these factors will be discussed later.

In the second case, the consumer may seek to estimate \( E(z) \) where \( E(z) \) is the expected difference in average expected characteristic utility
between the new alternative and the benchmark alternative (i.e., new minus benchmark). The consumer may not only require the estimated \( E(z) \) to be greater than zero before choosing, but require the estimated \( E(z) \) to be greater than some strictly positive quantity. In this case, the consumer feels that the loss in satisfaction resulting from mistakenly choosing the new alternative when it is, in fact, inferior to the benchmark alternative, is not worth the expected potential gains. If the consumer again allocates some fixed effort to evaluating the new alternative, a superior new alternative (i.e., superior to the benchmark) is more likely to be chosen when \( \frac{\text{VAR}(z)}{(E(z))^2} \) is relatively small. Hence, it is clear why companies seeking to promote superior new products stress the importance of the product field decision by claiming great variance in product characteristic utility within the field, large losses associated with choosing the wrong product in the field and the exceptionally attractive rewards associated with choosing the right product in the field. All of these claims promote the belief that a large confidence level should be assigned to the decision. This large level would lead to increased decision-making effort, fewer mistakes and hence a greater likelihood of choosing the new alternative.

In the final case, the consumer makes an unbiased evaluation of the new alternative. In other words, the consumer would choose the alternative if, after several characteristic comparisons, the consumer inferred that \( E(z) \) was greater than zero. In this case, it may be interesting to consider exactly why the new alternative was evaluated, i.e., what factors initiated the evaluation. For example, if during past recurring choices, the consumer chose different alternatives despite the fact that the elements of the alternative choice set remained constant, then a new entry into that alternative choice set would most likely be evaluated. One notable factor for causing continuous reevaluation of the identical alternative choice set would be changing alternative characteristic difference utilities. In other words, \( z_j \) may change over time for some subset of all possible characteristics \( j \). Of course, changes in the actual alternative characteristics could cause changes in \( z_j \). However, it is not necessary for actual alternative characteristic levels to change in order to observe a change in \( z_j \). Suppose the consumer becomes satiated at high levels of characteristic \( j \) and retains some of the benefits of characteristic \( j \) from the previous choice occurrence. Then \( U_j(x_j) \) may change while \( x_j \) does not. Further, \( z_j \) is equal to \( U_j(x_j) - U_j(y_j) \), so \( z_j \) could change because the function \( U_j(\cdot) \) changes over time even though \( x_j \) and \( y_j \) remain constant. A concrete example may lead to further insights. Suppose the consumer is selecting among Chicago restaurants. There are many Chicago restaurants and many restaurant characteristics (ability to impress guests, romantic quality, etc.). Let us consider the ratings of four Chicago restaurants and the associated characteristics.
restaurants on four restaurant characteristics as given in Table 2 (Kaplan, 1979).

Suppose the consumer has a linear utility function for food quality such as $2F + 3$ where $F$ represents the food quality rating in Table 2. Then the consumer’s characteristic utility for food for Fanny’s restaurant would be 15 while the characteristic utility for food for the Alouette restaurant would be 17. Hence, comparing the Fanny’s alternative with the Alouette alternative on the characteristic of food, the difference in characteristic utility (denoted $z_F$) would be $-2$ (i.e., $15 - 17$). However, if the decor utility function approaches some level of satiation (i.e., exhibits diminishing returns), then past experiences may effect future decisions. For example, if the characteristic utility function for decor is $7\sqrt{D} - 1$ and no previous choices have occurred, the difference in characteristic utility for decor (denoted $z_D$) would be $7\sqrt{4} - 7\sqrt{2}$ or $-4.0$ when comparing Waterfront with Topkapi. However, if the effects of decor were somehow accumulative and the pleasure deriving from experiencing a fine decor diminished approaching some level of satiation (or merely diminished in value as does $7\sqrt{D} - 1$), then the initial level of decor from past experience may be at some strictly positive level. In other words, the consumer may inventory the effects of certain characteristics. For example, suppose the initial level of decor when approaching the choice decision was 10. Then $z_D$ would equal $7\sqrt{4} + 10 - 7\sqrt{2} + 10$ or $-1.94$ when comparing Waterfront with Topkapi on decor. Note that $z_D$ changes because the square-root function exhibits diminishing returns (i.e., positive first derivative, negative second derivative). However, the value of $z_F$ would equal $2(6 + 10) - 2(7 + 10)$ or $-2$ as before when again comparing Fanny’s and Alouette on food. Therefore, when the utility function exhibits diminishing returns and some accumulation (inventory) effect occurs over time for some particular characteristics, the importance of those characteristics decreases over time. That is, the importance of these characteristics decreases as alternatives are selected which possess high levels of those characteristics. Here, $z_D$ decreased in absolute magnitude from 4.0 to 1.94. It is therefore clear that recurring changes in choice behavior can be provoked when these two conditions occur. In those situations, not only will the

<table>
<thead>
<tr>
<th>Restaurants</th>
<th>Food(F)</th>
<th>Service(S)</th>
<th>Decor(D)</th>
<th>Value(V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fanny’s</td>
<td>6</td>
<td>6</td>
<td>2.5</td>
<td>3</td>
</tr>
<tr>
<td>Waterfront</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Topkapi</td>
<td>6.5</td>
<td>5</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Alouette</td>
<td>7</td>
<td>6</td>
<td>3.5</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Cost_t_implications2.max
importance of these characteristics change over time, but the difficulty associated with the individual binary alternative choices will change over time. Finally, only a subset of all alternative characteristics may actually exhibit both this accumulation effect over time and this diminishing returns effect. Hence, new alternatives which are perceived as strong on characteristics having both of these properties will more likely be chosen than alternatives weak on these properties. This conclusion follows because consumers switching among alternatives for reasons of satiation would focus their attention (i.e., sample from) the subset of alternative characteristics exhibiting accumulation.

PERSUASION

There often exist situations when one decision-maker wishes to, with or without justification, influence the choice of a second decision-maker. When someone attempts to persuade someone else to choose one alternative over another, either at the time the choice is being made or at some preparatory stage, the attempt can be construed as a process of providing information. The information can be concrete, factual and very straightforward. For example, a restaurant ad may state that a particular restaurant is open until 10 p.m. The informed consumer now has more information on which to base the restaurant choice decision. This information may persuade the consumer to select this restaurant rather than rush to the previously selected restaurant which closes at 8 p.m. The information provided by the ad could also be abstract, subjective and less obvious. For example, a restaurant advertisement may picture a jovial family laughing and dining among other similar families. From this picture, the restaurant is attempting to vividly communicate a detailed message about the enjoyable atmosphere of their establishment. The message may subtly suggest numerous virtues of the restaurant from having clean floors to having sensual waitresses. In any case, whether information is hidden or not, the ad attempts to provide information which will tend to favorably influence the viewer of the ad toward choosing the advertised restaurant.

However, there is a difference between enlightenment and persuasion. With enlightenment, the consumer is given information to enable better decision making. With persuasion, information must be provided so as to have a desirable and predictable outcome on the consumer's decision process. Of course, persuasion may not be easy. First, consumers may realize that advertisers or other persuaders have the incentive to provide bias information. Second, consumers may have very little incentive to read or view the persuader's message and may thereby circumvent the persuasion effort completely.
Nevertheless, the methodology used to develop an index for measuring decision-making difficulty can also be used to develop a strategy for advertisers. Given the competitive structure of most markets, we seldom find companies whose brands dominate their competitors’ on every characteristic. Hence, it would not be prudent for most advertisers to reveal to the consumer every last detail of their brand. For example, persuading consumers to purchase a brand would seldom involve the revealing of the characteristics on which the brand is weak. The advertiser must be concerned with exactly how much information should be provided to the consumer through the ad message.

Intuitively, the optimal number of characteristics should provide the proper balance between revealing too much about the brand and not revealing enough. If the advertisement does not reveal a sufficient number of the brand’s attractive characteristics, the ad viewer may remain unconvinced and not purchase the brand. The ad would not be sufficiently revealing to entice the consumer. If the advertisement reveals too many of the brand’s characteristics, the viewer may not notice all the brand’s characteristics. In this case, there is a danger that the ad viewer may focus on some of the brand’s less attractive characteristics and ignore some of the more convincing characteristics. In this later situation, the consumer samples the ad for information on product characteristics and, in doing so, may infer the brand is less attractive than in the situation where the ad only contained the more convincing characteristics.

This analysis which views advertisements as populations of product characteristic information from which consumers sample information leads to some interesting results relating to the optimal amount of information an ad should contain (Shugan, 1979). However, solving this problem leads to deeper problems for theorists. For example, what is the correct definition of a product characteristic? How do we know if a product characteristic is actually two more detailed characteristics? And, do all alternatives in the choice set possess the same characteristics, only at different levels?

OTHER CHOICE PROBLEMS

The concepts of “thinking costs” and the “confusion index” are applicable to numerous areas not discussed here. Shugan (1980) derives the cost of using various simplifying decision rules. For example, the cost of a product comparison with the conjunctive rule is \((N/m - 1)/(1 - a)\) where \(N\) is the total number of characteristics used in the rule and \(m\) is the number of characteristics on which the product misses the minimum level. Shugan (1980) also derives distributions for mistakes.
when consumers reduce thinking effort. Tversky’s EBA (1972) is shown to be one special distribution.

The formulation of “thinking costs” for a utility maximizing model also implies methods by which thinking costs can be reduced. For example, thinking costs can be reduced through (1) memory, (2) collecting summary statistics, and (3) probabilistic sampling (Shugan, 1980). Remembering large \( z \) would reduce sampling effort. However, poor memory may encourage not considering all information and, hence, adoption of simplifying decision rules. Collecting summary statistics reduces thinking costs by lessening the number of necessary comparisons. However, summary statistics such as “overall quality ratings,” certifications or branding would be most useful when summarizing sets of characteristics where \( \text{VAR}(z)/[\text{E}(z)^2] \) is large. However, this type of summarization is only possible when consumers have common preferences because \( z \) is computed with the consumer’s utility function. Finally, by selecting only key attributes, the consumer can greatly reduce “thinking costs.” Hence a consumer buying a house for the first time may first seek out a book on “what to look for in a house” rather than information on particular houses. Experience may be related to knowing which characteristics to sample.

SUMMARY AND FUTURE RESEARCH

We have explored the concept of a thinking cost. First we reviewed the development of one measure for the difficulty of a decision using a utility maximizing procedure, i.e., the confusion index. We then explored its implications in the area of alternative deletion and found some interesting asymmetric error relations when decision-making effort is not properly allocated. Next we explored alternative addition and the complex problems of new alternative evaluation and choice over time. We also noted the effects of diminishing returns on “thinking costs.” The area of persuasion was then discussed. Our discussion led to some insights regarding optimal persuasive messages. Finally, some sundry choice problems and their relationships with thinking costs were explored.

However, there are many areas relating to decision-making effort which remain a total mystery. For example, do alternative characteristics naturally exhibit covariance? In competitive markets, do the distributions of product characteristics exhibit special properties such as negative covariance? What is the best way to summarize information to reduce decision-making effort when consumer tastes differ? What opportunities are forgone by the mistakes made as a consequence of reducing thinking costs, and how do we properly trade off decision-making effort with the
benefits of correct decisions? Clearly, despite the high activity in choice research, there are many problems relating to decision-making which remain unsolved.

ACKNOWLEDGMENT

Prepared for the Interdisciplinary Choice Theory Conference, University of Washington, Seattle Washington

REFERENCES


Consumer Reports, 1980, 45(1), 8-11.


