The Role of Consumers’ Intuitions in Inference Making

SUSAN M. BRONIARCZYK
JOSEPH W. ALBA*

Traditional explorations of inference making have examined consumers’ reactions to product descriptions that lack information about salient attributes. Such studies frequently report systematically lower evaluations of incompletely described alternatives along with a generally low incidence of unprompted attribute-to-attribute inference. We argue that the nature and likelihood of an inference are dependent on the cues available at the time of decision making, and that some cues may exert a disproportionate influence on inference behavior. In several experiments in which subjects were presented with competing cues that implied different values of a missing attribute, we show that intuitive beliefs about the relationships between attributes are perceived as a particularly reliable basis for interattribute inference. Strong beliefs appear capable of superseding other compelling cues and may induce consumers to generate inferences spontaneously.

Inference making has played an enduring role in the explanation and prediction of consumer behavior, because it is widely recognized as a source of product beliefs and an attractive substitute for search and direct product experience. Anecdotally, inference making appears to be a central managerial consideration in many advertising, branding, packaging, pricing, product design, and retailing decisions. However, much of our formal understanding of consumer inference is derived from research conducted to test assumptions—typically that individuals possess full information or ignore missing information—associated with particular market research methods or models of human judgment. This research has properly demonstrated that judgment processes may involve inference making in addition to, or apart from, the strict adding or averaging of explicit information (see Cohen, Miniard, and Dickson 1980; Johnson and Levin 1985; Levin, Johnson, and Faraone 1984), and that inference making can exert a significant but undetermined role in the measurement of attribute importance weights and forecasts of consumer preference (cf. Huber and McCann 1982; Meyer 1981).

Surprisingly, however, this same research has reported only sparse levels of spontaneous interattribute inference (see Huber and McCann 1982; Lim, Olshavsky, and Kim 1988). Interattribute inference refers to the case in which the value of a missing product attribute is inferred from another attribute of the same brand—such as when the quality of a brand is inferred from its price. In place of such reasoning, subjects in these studies appear to compensate for a missing attribute by computing the average value of that attribute across all brands. This inferred value may then be discounted to adjust for uncertainty prior to being incorporated into the evaluation of the brand (Jaccard and Wood 1988; Meyer 1981; Ross and Creyer 1992). However, recent evidence suggests that even such conservative behavior may be rare. Simmons and Lynch (1991) employed a direct process measure and found a generally low incidence of inference making. They concluded that consumers are more likely to refrain from inference making and treat missing information as a negative cue that they then integrate into the overall evaluation of the object.

Such findings may be driven in part by method. Most of the studies that have failed to find evidence of spontaneous inference have employed a procedure involving the serial presentation and evaluation of numerous product alternatives described along two or three common dimensions. Embedded in the set are some incompletely described alternatives that lack information about one or two of the dimensions. Inference is assessed by comparing the evaluations of the complete and incomplete profiles. For reasons described later, this paradigm may not be sensitive to the full range of product inferences or decision situations. In fact, several studies of advertising and persuasion that have departed from the traditional paradigm portray consumers as willing elaborators (Kardes 1988; Pechmann 1992). At a minimum, these mixed findings suggest that inference
making is context sensitive. Thus, it behooves consumer researchers to understand its determinants. Fortunately, prior research has shed considerable light on the factors that affect the sheer frequency of inference making. Specifically, the likelihood of spontaneous inference making has been shown to vary as a function of the need for inference, knowledge of the domain, involvement in the task, the amount of contextual information supporting the inference, the nature of the task, and practice with the inference rule (Alba and Hutchinson 1987; Harris and Monaco 1978; Hastie and Park 1986; McKoon and Ratcliff 1992; Sanbonmatsu, Kardes, and Herr 1992; Smith 1984; Stayman and Kardes 1992; Uleman 1987).

A second, less understood aspect of inference making concerns the informational basis on which consumers are willing to form an inference. Even highly motivated consumers should not be expected to generate or use inferences unless a plausible basis for inference exists. The accessibility-diagnosticity framework of Fieldman and Lynch (1988) implies that information will serve as a basis for inference making only insofar as it is available and is perceived to be a reliable predictor of the missing value.\(^1\) Many bases for inference exist, but few have been explored. Moreover, multiple bases may exist within any situation, but little is known about the dominance relations among them. However, two notable exceptions exist.

Dick, Chakravarti, and Biehal (1990) compared the relative influence of evaluative consistency and probabilistic consistency as bases for inference. Evaluative consistency refers to a reasonable strategy by which the value of a missing attribute is inferred to conform to the overall evaluation of the brand (see Fishbein and Ajzen 1975). Thus, a missing attribute will be assigned a favorable value if the overall evaluation of the brand is favorable. Probabilistic consistency refers to another plausible strategy by which consumers infer a missing attribute on the basis of its perceived causal or correlational relationship with a known attribute (e.g., inferring quality from price). In their experiment, probabilistic consistency was operationalized as an empirical interattribute correlation. Specifically, the critical attribute that was absent from the incompletely described alternatives was perfectly correlated with a different attribute in the completely described alternatives. Dick et al. (1990) found not only that subjects were willing to rely on the interattribute correlation, but also that this was preferred to evaluative consistency as a basis for inference.

The second exception is provided by Ross and Creyer (1992). They demonstrated that the amount of other-brand variance on the to-be-inferred attribute serves as an important determinant of inference making. When variance among known brands is low (i.e., all brands perform at a near-equal level), the same level may be inferred to exist in brands lacking explicit information about the attribute. Thus, low other-brand variance provides a compelling basis for inference.

Both studies suggest that data-based cues are perceived as valid predictors of missing information. In the Dick et al. (1990) study, the dominant cue (interattribute correlation) was derived from the actual relationships among the attributes in the stimulus descriptions. In the Ross and Creyer (1992) study, the dominant cue (other-brand variance) also was wholly contained in the brand-by-attribute matrices that served as the stimuli.

**THEORY VERSUS DATA**

Although it is comforting that consumers may be sensitive to the empirical relationships that exist among attributes and brands, there remains a question regarding the extent to which data-based cues will influence inference making. The question is prompted by two observations. First, other, more persuasive inference cues may overshadow the effects of the data. Second, if data-based cues are truly persuasive, they should prompt spontaneous inference making. In the Ross and Creyer (1992) study, spontaneity was not of focal interest and therefore was not measured. In the Dick et al. (1990) study, spontaneity was measured, but few instances of spontaneous inference were observed.

In this regard, it is worth noting an additional aspect of the Dick et al. (1990) study. The stimulus consisted of a set of 35-mm cameras in which the incomplete brands lacked information about lens sharpness. In the completely described brands, lens sharpness scores were given and were shown to be perfectly correlated with shutter speed. Such a strong correlational relationship can provide a compelling probabilistic basis for inference making (as confirmed by the behavior of subjects in the prompted-inference condition). However, relationships that are strictly correlational may not satisfy a consumer’s explanatory criteria and therefore may be perceived as having limited generalizability. Einhorn and Hogarth (1986) argue that, if a causal schema or theory cannot be constructed to bridge two events, the events may be perceived as causally unrelated, regardless of their rate of cooccurrence. Similarly, Ajzen (1977) has demonstrated that the utilization of base rate data in a prediction task depends largely on the extent to which the base rate is perceived to be causally related to the predicted event. Thus, the use of one variable to predict another may depend not only on the strength of their empirical correlation but also on the nature of their intuitive relationship (Broniarczyk and Alba 1994a; Jennings, Amabile, and Ross 1982).

We suggest that similar effects may characterize interattribute inference making. A particular missing value may appear “probable” from a statistical per-
spective but not from an intuitive perspective. In the
following experiments we investigate the effects of the-
ory-based relationships vis-à-vis empirical relationships.
We use the term theory to refer to consumers' intuitive
beliefs about the likely relationships among attributes.
These implicit theories reflect consumers' understanding
of causality and need not be logically correct, empir-
ically accurate, or even the result of direct experience
(see Nisbett and Wilson 1977). For example, consumers
may believe that price is predictive of quality on the
basis of the theory that higher prices are driven by better
component parts or workmanship. There is ample ev-
eidence that consumers generate inferences on the basis
of their general beliefs about the world (see, e.g., Alba
and Hutchinson 1987; Sujan and Dekleva 1987). How-
ever, the relative potency of such beliefs has received
less attention, particularly in the domain of attribute-
to-attribute inference. The present experiments exam-
ine the influence of such beliefs in the face of contra-
dictory data. Specifically, we construct situations in
which the inference favored by intuition conflicts with
the inference favored by the empirical relationships that
exist among brands and attributes; we then measure the
extent to which subjects' inferences are theory-driven
versus data-driven.

Although our primary goal is to examine the relative
influence of theory and other potential inference cues, we
also examine the effects of theory on spontaneity.
We argue that spontaneity requires not only a favorable
method of measurement but also the presence of suf-
ficiently persuasive inference cues. Inference making is
an inherently risky activity, because all inferences are
probabilistic by nature. Consumers may incorporate
probabilistic information into a decision only when they
are reasonably confident that they have elaborated
properly and that elaboration is necessary for proper
decision making.

In experiments 1 and 2, we pit theory-based attribute
relations against data-based attribute correlations. In
experiments 3 and 3A, we examine the implications of
theory for spontaneous inference. In experiment 4, we
pit theory against other-brand variance in the to-be-
inferred attribute. In experiment 5, we examine the in-
fluence of theory from a more pragmatic perspective.
Because our primary interest is in studying the relative
influence of competing inference cues, we borrow the
paradigm developed by Dick et al. (1990). This para-
digm relies jointly on the nature of the stimulus design
and the decision outcomes to identify dominance re-
lations among inference cues.

Stimulus Selection

Before describing the experiments, it is necessary to
explain our choice of stimulus materials. Throughout
this research, we examine the willingness of subjects to
infer a missing attribute on the basis of its intuitive
(i.e., theory-based) relationship to a known attribute.

In virtually every experiment, the attribute pair chosen
to represent an intuitive relationship consisted of repair
record (or durability) and warranty. Warranty served
as the known attribute; repair record (or durability)
served as the missing attribute. We chose this pair for
two primary reasons. The first relates to external valid-
ity. Warranty and repair record possess an obvious intu-
itive connection (see Boulding and Kirmani 1993;
Shimp and Bearden 1982). It is easy to construct an
explanation of why these two dimensions should be
positively correlated (e.g., manufacturers of low-quality
goods should seek to minimize their liability). However,
these two attributes are not deterministically related,
and an inverse relationship is also easy to justify and
document (e.g., warranties represent a cost to the man-
ufacturer and may be most appealing to manufacturers
who need to overcome a reputation for low quality, as
in the case of American automobile firms). Conse-
quently, any warranty–repair record inference is prob-
abilistic, and the risks of an erroneous inference may
be considerable.

Also related to external validity is our desire to in-
vestigate attributes that are likely to be the target of
inference making in real product decisions. An infer-
ence is most likely to be generated by consumers when
they wish to reduce uncertainty about an attribute that
is important but difficult to assess through search or
short-term experience. Repair record fits this descrip-
tion. On the other hand, the basis for the inference
should be concrete and available. Warranty information
is unambiguous and available for those who search
for it.

Two pretests were conducted to test the validity of
our choice of attributes. The first pretest examined the
relative perceived strength of the warranty–repair record
relationship vis-à-vis other attribute pairs within the
product class of 35-mm cameras. A group of 26 subjects
was asked to rate the strength of relationship among 10
attributes (or 45 unique attribute pairs) on a nine-point
scale that ranged from "very much" to "not at all." Results
showed that the warranty–repair record pair re-
ceived the highest rating of 7.81, whereas the other pairs
clustered near the midpoint of the scale. In particular,
the warranty–repair record rating was significantly
higher than the 5.77 rating assigned to the shutter
speed–lens sharpness pair ($t(1,25) = 4.47, p < .001$),
which had served as the correlated pair in the Dick
et al. (1990) study and was used as a control in the present
research.

The second pretest examined the strength of the war-
 ranty–repair record relationship vis-à-vis evaluative consisten-
ty. Dick et al. (1990) found that a correlation-
based cue is perceived as more reliable than evaluative
consistency as a basis for inference. Because we wish to
contrast correlation with theory, it was desirable to show
that theory also dominates evaluative consistency. To
do so, 30 subjects were presented a brand-by-attribute
matrix describing four fictitious brands of 35-mm cam-
TABLE 1
STIMULI FOR EXPERIMENT 1

<table>
<thead>
<tr>
<th>Attributes</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expansion Flexibility</td>
<td>Very Flexible</td>
<td>Not Flexible</td>
<td>Quite Flexible</td>
<td>Somewhat Flexible</td>
</tr>
<tr>
<td>Convenience of Use</td>
<td>Somewhat Conven</td>
<td>Quite Conven</td>
<td>Very Conven</td>
<td>Not Conven</td>
</tr>
<tr>
<td>Warranty</td>
<td>3 Months</td>
<td>24 Months</td>
<td>6 Months</td>
<td>12 Months</td>
</tr>
<tr>
<td>Maximum Shutter Speed</td>
<td>1000</td>
<td>500</td>
<td>250</td>
<td>2000</td>
</tr>
<tr>
<td>Repair Record/Lens Sharpness</td>
<td>9</td>
<td>4</td>
<td>1.5</td>
<td>19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attributes</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expansion Flexibility</td>
<td>Somewhat Flexible</td>
<td>Quite Flexible</td>
<td>Very Flexible</td>
<td>Not Flexible</td>
</tr>
<tr>
<td>Convenience of Use</td>
<td>Not Conven</td>
<td>Very Conven</td>
<td>Somewhat Conven</td>
<td>Quite Conven</td>
</tr>
<tr>
<td>Warranty</td>
<td>12 Months</td>
<td>6 Months</td>
<td>3 Months</td>
<td>24 Months</td>
</tr>
<tr>
<td>Maximum Shutter Speed</td>
<td>2000</td>
<td>250</td>
<td>1000</td>
<td>500</td>
</tr>
<tr>
<td>Repair Record/Lens Sharpness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE.—In both phases of the experiment, the fifth attribute was lens sharpness in the control condition and repair record in the theory condition. Across conditions, only the label of the fifth attribute differed. All other attribute labels and all attribute values were held constant.

eras along four dimensions (cf. Dick et al. 1990). One of the brands was clearly superior to the others on three of the four dimensions and also possessed the second highest warranty. The brand with the highest warranty was somewhat average on the other dimensions taken together. No correlations existed among the attributes.

Subjects were asked to weight the attributes equally and to identify the best alternative. Afterward, they were presented with the same information, with either repair record index or lens sharpness index listed as the fifth attribute. However, the values of the fifth attribute were unspecified. Instead, subjects were asked to rate each brand on this missing dimension on a 20-point scale. Responses were scored simply in terms of which brand was rated most highly on the missing dimension.

Results showed that all 15 subjects who were asked to infer lens sharpness assigned the highest rating to the brand that was superior overall, 11 of whom explicitly referred to its overall superiority as the reason for their judgment. Among subjects who inferred repair record, 10 of 15 assigned the highest rating to the brand with the highest warranty, nine of whom explicitly referred to warranty as their rationale. The difference between the groups was significant ($X^2 = 10.79, p < .01$). These results indicate that an intuitive relationship may dominate evaluative consistency when the two are pitted against each other (see also Peabody 1967).

EXPERIMENT 1

In this experiment we perform a stronger test of the influence of theory-based relationships by examining consumers’ reliance on intuitive relationships in the face of contradictory data. The general method adheres closely to the paradigm developed by Dick et al. (1990). During the study phase, subjects viewed a set of brand descriptions that portrayed particular correlations among the attributes. During the subsequent test phase, a second set of brands was described along the same dimensions; however, values for one of the dimensions were missing. Subjects were asked to infer these values. The test brands were designed to reveal the basis for the inferences. The major departure from Dick et al. (1990) involved a between-subjects manipulation of the intuitive relationship between the missing attribute and the presented attributes. In the theory condition, the missing attribute (repair record) was intuitively related to one of the explicitly presented attributes (warranty). In the control condition, the missing attribute (lens sharpness) possessed no compelling intuitive relationship with any of the explicitly presented attributes.

Method

Study Phase. A total of 41 subjects were presented with the brand-by-attribute matrix depicted in Table 1 in a gift-giving context and were asked to choose the best brand for a friend, guided by the assumption that the friend valued the attributes equally. The attributes, their specific values, and their spatial proximity were chosen to maximize similarity to the stimuli used by
Dick et al. (1990). The first two attributes, expansion flexibility and convenience of use, were described on four levels that ranged from "not flexible" to "very flexible" and "not convenient" to "very convenient," respectively. The warranty dimension was described in months, ranging from three to 24 months. Maximum shutter speed had four levels, which ranged from 1/250th to 1/2,000th of a second. The fifth attribute was manipulated between subjects. For the 21 subjects in the control condition, the attribute was lens sharpness. For the 20 subjects in the theory condition, the fifth attribute was repair record. Regardless of condition, the numerical values of the fifth attribute were identical and were expressed on a 20-point scale. Note that a perfect correlation exists between the manipulated attribute and the shutter speed values immediately above them. Note also that no correlation existed between the manipulated attribute and warranty or between shutter speed and warranty. The remaining attribute values were assigned in such a way that none of the alternatives would appear to have an overwhelming advantage. Furthermore, the attribute value assignment resulted in a perfect negative correlation between convenience and the manipulated attribute, which provided an additional opportunity for data-driven inference.3

Test Phase. After each subject chose the best brand from among brands A–D, the brand descriptions were removed. The cover story then told subjects to assume that all four brands were beyond their price range, but that four additional brands (E–H) were affordable. The only drawback was that these latter brands lacked information about the manipulated attribute (i.e., lens sharpness vs. repair record). Brands E–H were then presented and subjects were asked to infer the missing values for each brand with the use of the 20-point scale. Afterward, rationales for the responses were elicited.4

Inspection of brands A–H reveals that each brand in the second set is identical to one of the brands in the first set with the exception of the missing attribute. Thus, subjects were provided with several bases for inference. All subjects had viewed a perfect positive correlation between shutter speed and the missing attribute (as well as a perfect negative correlation between convenience and the missing attribute). Inferences based on interattribute correlation would provide brand E with the highest rating, because the positively correlated attribute of shutter speed was highest for this brand. (In addition, the negatively correlated convenience attribute was lowest for this brand.) Also, with the assumption of reasonable memory for brands A–D, exemplar-based inferences were possible. Exemplar-based inferences favor brand E, inasmuch as the brand corresponding to brand E in the initial set (brand D) possessed the highest value of the missing attribute. Thus, both interattribute correlation and exemplar-based reasoning favor brand E. Theory alone favors brand H for those subjects who were asked to infer repair records, because brand H possessed the longest warranty.

Results and Discussion

The brand given the highest rating on the missing attribute dimension by each subject served as the dependent measure. All but six subjects, four of whom were in the control condition, rated either brand E or brand H most highly. These subjects were deleted from the analysis. The remaining subjects confirmed our hypothesis. In the control condition, 14 of 17 subjects rated brand E most highly. Of these, seven subjects referred to exemplar-based reasoning, and six referred to the speed-sharpness correlation in brands A–D. On the other hand, 11 of 18 subjects in the theory condition rated brand H most highly. Of these, 10 referred to warranty and one alluded to exemplar-based reasoning. The difference between conditions in terms of which brand was rated most highly was significant ($X^2 = 6.00, p < .02$).

Thus, when no intuitively related predictor conflicted with the data (control condition), subjects appeared to be attuned to, and influenced by, the regularities in the attribute correlation structure. By comparison, reliance on the data was much lower when the data were in conflict with theory. Indeed, the tendency was to rely on theory even when unambiguous evidence argued against it. Subjects in the theory condition shifted significantly toward brand H on the basis of its warranty despite (i) its low shutter speed, which was perfectly correlated with repair record during the study phase, and (ii) empirical disconfirmation of a relationship between repair record and warranty during the study phase.

The generalizability of these results may be questioned, because they are somewhat distorted by the high use of exemplar-based reasoning. It was not our intent to investigate exemplar-based inference, and we were surprised by the reliable memory subjects exhibited for brands A–D. However, our basic point is not compromised. Indeed, the reliance on theory by subjects in the theory condition is made more compelling by the fact that the theory was opposed not only by the interattribute correlations but also by some apparently salient exemplar information.

Nonetheless, the unexpected emergence of the exemplar strategy prompted us to conduct a follow-up study that eliminated the possibility of exemplar-based processing. Specifically, the stimuli were altered in such a way that the duplication among brands between the study and test phases was eliminated, but the empirical
relationships among warranty, shutter speed, and sharpness/repair were essentially preserved. The procedure was unchanged in all other respects. Results from subjects who chose brands directly pertaining to the theory/data comparison strongly supported our hypothesis. In the theory condition, eight subjects chose the brand with the highest warranty, but only two subjects chose the brand predicted by interattribute correlation (i.e., the brand with the highest shutter speed); in the control condition, 11 subjects chose the brand predicted by interattribute correlation, and no subject chose the brand with the highest warranty ($X^2 = 9.17$, $p < .01$). Thus, the results seem generalizable in the confines of the basic paradigm.

Subjects' apparent readiness to infer one attribute from an intuitively related attribute, even when the empirical relationship between them has been disconfirmed, is a testament to the strength of theory. A logical follow-up question concerns whether the perceived validity of an intuitive relationship leads to unprompted inference making. Before turning to this issue, however, it is necessary to address a second threat to the generalizability of experiment 1. Specifically, the correlational relationships among the attributes of the study phase brands (brands A–D) were unambiguous but appeared over a small set of brands. Thus, subjects may not have noticed the perfect correlation between lens sharpness (or repair record) and shutter speed or the lack of correlation between repair record and warranty. In addition, the biasing effects of prior beliefs may have caused subjects in the theory condition to erroneously perceive a correlation between warranty and repair record (i.e., an illusionary correlation; Chapman and Chapman 1969). Finally, because brands E–H were described as lower in price than brands A–D, subjects may have refrained from generalizing from the study set to the test set, irrespective of how well they learned the correlational relationships. Experiment 2 addresses these issues in the context of prompted inference, and follow-up experiments examine spontaneity.3

### EXPERIMENT 2

Instead of examining four brands described across five attributes during the study phase, subjects in this experiment viewed 25 brands described along three dimensions. Two of the dimensions, durability and shutter speed, were intuitively unrelated but were highly correlated. Thus, subjects had ample opportunity to learn a clear empirical relationship. In another departure from the previous experiments, no mention of the intuitively related predictor (warranty) was made during the study phase, which eliminated the possibility of illusionary correlation. Finally, manipulation of theory occurred solely during the test phase. All subjects were asked to infer durability, but only the subjects in the theory condition were provided with warranty information at the time of prediction.

### Method

Prior to viewing the brands, subjects were informed that they would see information on 25 randomly chosen 35-mm cameras that were representative of the entire class. Three dimensions of information were then described in terms of their meaning and units of measurement. Durability was defined as overall durability and was expressed on a 100-point scale. Weight was expressed in ounces, and it was noted that lighter cameras are easier to carry. Maximum shutter speeds were noted as ranging from the slowest speed of 1/500th of a second to the fastest speed of 1/4000th of a second. Subjects were told to assume that they were evaluating the cameras for a friend who valued the three attributes equally. To prevent cursory processing of the data and to enhance learning of the empirical relationships, subjects were required to evaluate each alternative on a seven-point scale.

In the actual stimulus set, the brands were labeled A–Y, and each was accompanied by values for durability, weight, and shutter speed. Across the brands, durability ranged from 14 to 98 on the 100-point scale. Weight ranged from 20 to 35 ounces, and shutter speed assumed the range of values listed above. The correlation between durability and speed was .93; the durability-weight correlation and weight-speed correlation were both near zero.

After the ratings were completed, the test phase began. Subjects were informed that they would view 15 additional brands similar to the previous 25, but that these brands lacked durability scores. The task was to estimate durability for each new brand on the basis of shutter speed, weight, and two additional attributes. For the 19 subjects in the control condition, the additional attributes were picture sharpness and expansion flexibility; for the 18 subjects in the theory condition, the additional attributes were warranty and expansion flexibility. The inclusion of expansion flexibility in both conditions was done to reduce and test for demand effects.

Regardless of the condition, all subjects viewed the same data. The only difference between the theory and control conditions was in the label given to one of the four attributes used to describe the test brands. Across the 15 test brands, weight and speed assumed the same range of values as in the learning phase. The new attributes were described on a scale that ranged from “poor” to “excellent.” (Verbal descriptors were used

---

3In fact, an experiment was conducted to assess the extent to which theory leads to spontaneous inference in the context of the stimulus design of experiment 1. As predicted, subjects in the theory condition generated spontaneous inferences and did so at a rate significantly higher than subjects in the control condition. We do not present the specifics of this experiment for purposes of brevity and, moreover, because the results and conclusions are largely contained in subsequently reported experiments. Details are available on request.
for the new attributes so that warranty, sharpness, and flexibility could be expressed on an identical scale.) The intercorrelations among the four attributes all were near zero. Subjects made their durability ratings by means of the same 100-point scale used to describe the initial set of 25 brands.

Results and Discussion

Regression analyses were performed to estimate the weights placed by each subject on each of the four given attributes when predicting durability of the test brands. The mean standardized weights are presented in Table 2. (For the purpose of analysis these beta weights were transformed to Fisher's Z-scores, and contrasts were performed within conditions.) It is evident that subjects in the control condition were most influenced by the data presented during the learning phase. The attribute receiving the most weight was shutter speed, which was highly correlated with durability during the learning phase. The weight placed on speed was significantly higher than the weight placed on any other attribute ($p's < .01$). The heaviness of the camera had virtually no influence on inference, which is appropriate given its lack of empirical correlation with durability. The new attributes, sharpness and flexibility, received moderately low and equal weights ($F < 1$).

A different pattern is observed in the theory condition, where warranty was most influential. Although not statistically significant, subjects placed a higher weight on warranty than on the empirically correlated attribute of shutter speed (.63 vs. .42, respectively, $F_{1,16} = 1.66, p > .10$). Both warranty and shutter speed received significantly higher weight than flexibility or heaviness of the camera ($p's < .03$). Thus, theory subjects were influenced by speed to a nontrivial extent and were virtually uninfluenced by warranty or flexibility. The low weight received by flexibility in both conditions assures that the two additional attributes introduced during the test phase did not “demand” subjects' attention.

The means suggest that the weight placed on warranty by subjects in the theory condition was as high as the weight placed on shutter speed by subjects in the control condition (both $\beta's = .63$) and was nearly twice as high as the weight placed on sharpness by subjects in the control condition (.63 vs. .32, respectively). To make statistical comparisons across conditions, subjects were classified according to whether they placed a higher weight on shutter speed or on the manipulated attribute. Results showed that only 11 percent of the subjects in the control condition placed a higher weight on sharpness than on shutter speed (the empirically correlated attribute); in contrast, 67 percent of the subjects in the theory condition placed a higher weight on warranty (the theory-based attribute) than on shutter speed ($\chi^2 = 9.46, p < .003$).

Thus, despite extensive evidence for the reliability of shutter speed as a predictor of durability, subjects exposed to warranty information relied extensively on warranty—an attribute for which they had no direct evidence. In this regard, it is important to note that the learning-phase instructions explicitly alerted subjects to the fact that the initial set of 25 brands was representative of the entire product class. In purely predictive terms, therefore, it would be difficult to improve performance by incorporating any variable other than shutter speed into the judgment. Insofar as subjects would attempt to improve predictive accuracy beyond the level provided by shutter speed, it is not surprising that warranty would receive attention, given that there is little basis for relying on weight or flexibility. However, the magnitude of the weight placed on warranty testifies to the power of theory relative to data.

EXPERIMENT 3

Given that theory can exert a strong influence on the content of an inference, the question of spontaneity merits attention. Experiment 3 adopts elements of the previous experiments but employs a different dependent measure to detect the occurrence of spontaneous inference making. As in experiment 1, the independent variable was manipulated in the study phase; in all other respects, the study phase was identical to that of experiment 2. In the test phase, however, a choice task was substituted for the prompted inference task employed in experiments 1 and 2. Following the logic of Dick et al. (1990), the choice alternatives were constructed in such a way that the choice outcomes could indicate the occurrence of spontaneous theory-based inference.

Method

As in experiment 2, all subjects were provided with the gift-giving scenario that required them to inspect 25 brands of cameras. These brands were said to constitute the entire set of brands on the market. The friend for whom the gift was intended was a professional nature photographer who equally valued only three attributes in a camera: (a) low weight, (b) fast shutter speed, and
(c) high picture sharpness or high durability. The latter attribute was manipulated between subjects, as in experiment 1. Picture sharpness served as the third attribute in the control condition and durability served as the third attribute in the theory condition. Except for the label assigned to the third attribute, the stimulus set was identical to the one described in experiment 2. Thus, among the three attributes, the only nonzero correlation existed between shutter speed and the manipulated attribute (sharpness or durability), which were correlated at .93.

After rating the 25 brands on a seven-point scale by means of an equal-weights decision rule, subjects were informed via the gift-giving scenario that two promising new brands had just appeared on the market. However, because of the recent introduction of these brands, it was not possible to assemble all the important information. The only source of information was in the materials that accompanied the cameras. All subjects were presented with descriptions of the two brands, which were purportedly taken from these materials, and were asked to make a choice. Subjects were reminded to choose on the basis of their friend’s preferences rather than their own and to assume that their friend definitely would prefer one brand over the other.

The critical attributes of the test brands were embedded among three additional distractor features (see Table 3). Of the attributes important to the hypothetical friend, both cameras weighed 26 ounces. Brand X possessed a maximum shutter speed of 1/2,000th of a second; brand Y possessed a superior speed of 1/4,000th second. No information was provided regarding either of the manipulated attributes (durability or sharpness). However, brand X possessed a 24-month warranty whereas brand Y possessed a six-month warranty.

On the basis only of explicit attributes that were deemed important in the instructions, brand Y was a dominating alternative (i.e., it was superior on shutter speed and equivalent on weight). Brand Y also would dominate if its missing attribute were inferred from shutter speed, because shutter speed favored brand Y and was highly correlated with the missing attribute. The only rationale for preferring brand X would occur among subjects in the theory condition who not only ignored the speed-durability correlation but also inferred a durability advantage for brand X (on the basis of warranty) that exceeds the speed advantage of brand Y. Note that the designs of brands X and Y provide a conservative test of spontaneity. Shutter speed not only was highly predictive of durability but also was deemed to be an important choice attribute itself. In addition, subjects may infer durability from warranty but still opt for brand Y, because it dominates on the nonprobabilistic attributes. Thus, the proportion of subjects in the theory condition who choose brand X may underestimate the proportion who generated a spontaneous inference on the basis of theory.

Results and Discussion

Of the 10 subjects in the control condition, nine selected brand Y, as expected. Of the 11 subjects in the theory condition, six chose brand X. The difference between conditions was reliable ($X^2 = 3.60, p < .06$). Consistent with prior research, none of the subjects in the control condition referred to the speed-sharpness correlation. These subjects apparently were content to choose the alternative that dominated on the basis of the presented attributes, although it is not possible to rule out the alternative that the postchoice rationales were insufficiently sensitive to tap inference making. On the other hand, the six subjects who chose brand X in the theory condition all referred to warranty as the basis for their choice. Given that warranty was deemed unimportant in choice and was not mentioned by any of the control subjects, these responses are supportive of spontaneous inferences. Furthermore, half of these subjects explicitly stated that they inferred durability from warranty.

These results are consistent with those of the preceding experiments. It is interesting that the proportion of subjects in the theory condition who appeared to generate a spontaneous inference compares favorably with the highest levels of spontaneous inference reported in the literature. The high-water mark appears to have been reached by Levin et al. (1984), who employed the traditional judgment paradigm and found evidence for inference among one-third of their subjects in one of their conditions. This result was obtained when subjects had the opportunity to infer quality from price and viewed profiles in which quality and price were perfectly correlated. That is, both theory and data supported the inference. In the present experiment, a higher level of spontaneous inference was obtained when (i) an intuitive but empirically disconfirmed relationship supported the inference and (ii) a different empirical relationship predicted an opposite inference. In addition, our measure is conservative, as it may not detect risk-averse subjects in the theory condition who may have made an inference but disregarded it in favor of the explicit attributes favoring brand Y.

**EXPERIMENT 3A**

A replication of experiment 3 was performed to ensure generalizability and to test the assertion that spon-
taneous inference is highly sensitive to the measurement context. In particular, the traditional paradigm, which asks subjects to form individual evaluations of numerous product profiles, may reduce the motivation for inference making (Huber and McCann 1982; but see Simmons and Lynch 1991). Less effort is required if one attends only to explicit information, especially if multiple judgments are solicited. Similarly, the very nature of the decision task (i.e., product judgment rather than product choice) may suppress inference making. Because judgment tasks foster little sense of correct versus incorrect responses, the perceived need to arrive at optimal decisions or to discriminate among alternatives may be relatively low.

Thus, experiment 3A examined an intuitive relationship that has been used in previous research but employed a choice task in place of a judgment task. Carpet cleaners served as the product class, and the intuitively related attributes were tank size and area cleaned per refill. These attributes are related almost by necessity and therefore should prompt high levels of spontaneous inference. However, in previous research employing the traditional paradigm, these same attributes have been reported to produce low levels of spontaneous inference (Simmons and Lynch 1991).

Method

The procedure and stimulus structure were identical to those used in experiment 3. Only the attribute names and the scales used to express the values of each attribute were changed. All of the correlational relationships were preserved exactly. Specifically, the attribute labels of weight, shutter speed, and sharpness or durability from experiment 3 were replaced by drying time, warranty, and versatility or area cleaned per refill, respectively. (Where necessary, the attribute values were rescaled for realism.) Thus, unlike the previous experiments, warranty was intuitively unrelated to the manipulated attribute but was highly correlated (.93) with it.

As in experiment 3, subjects first rated the 25 brands on a seven-point scale by means of an equal-weights decision rule. In the test phase, all subjects chose between brands X and Y, which were described in terms of some distractor features and two out of the three important features. Specifically, both brands possessed identical drying times, but brand Y had a superior warranty (24 vs. 12 months). No information was provided about either versatility (control condition) or area cleaned per refill (theory condition). However, brand X possessed a larger water tank (4 gallons vs. 1 gallon).

SIMMONS ET AL. 31

Simply on the basis of the explicit attributes or on the correlational relationship between warranty and the missing attribute, brand Y dominated brand X in both experimental conditions. That is, brand Y should be chosen if one attends only to the explicit information deemed important, because it had a superior warranty and equivalent drying time. Likewise, brand Y should be chosen by subjects who make correlation-based inferences, because it had better performance on the correlated attribute of warranty. Theory alone favors brand X for those noncontrol subjects who opt to infer area cleaned from its intuitive relationship to tank size.

Results and Discussion

Two subjects failed to follow instructions and were dropped from the analysis. The following results indicate that their absence had no effect on the outcome. As expected, a preference for brand Y was exhibited in the control condition, wherein six of seven subjects chose the brand that dominated on the explicit attributes. Also as expected, subjects in the theory condition preferred brand X nearly unanimously (eight out of nine). The difference between conditions was significant ($X^2 = 6.63, p < .01$). In addition, the rationales were unambiguous. Of the eight subjects who chose brand X, all mentioned tank size as their rationale, with seven explicitly indicating that they inferred cleaning area from tank size.

One is tempted to argue that these results are unsurprising given the nature of the attributes. However, the results must be viewed in light of prior research. As noted, Levin et al. (1984) employed attributes that were strongly related intuitively, empirically, and ecologically but produced relatively low rates of inference. Moreover, Simmons and Lynch (1991) employed intuitive relationships that were virtually identical to the one used here, sometimes supplementing them with perfect empirical correlations; yet they also uncovered a paucity of spontaneous inference. However, both studies employed the traditional judgment paradigm. Combined with the fact that inference making is inherently risky and often requires some degree of effort, judgment tasks may provide little motivation for generating or incorporating probabilistic beliefs into one's evaluations. On the other hand, choice tasks, such as the one employed here, imply the existence of a best alternative and may motivate consumers to take the risky step of incorporating probabilistic information into their deliberations. The variance in outcomes observed across experiments argues that the extent to which consumers generate and use inferences may vary widely as a function of subtle experimental factors, as well as of product class.

The remaining experiments are redirected at the issue of the dominance relations among competing inference cues. The preceding experiments contrasted theory to interattribute correlation. As noted at the outset, Ross and Creyer (1992) have shown that low other-brand variance on the to-be-inferred attribute can also serve as a highly influential data-based cue in the context of inference. In the next experiment, this cue is contrasted to theory.

**EXPERIMENT 4**

Ross and Creyer (1992) clearly demonstrated that, when variance among known brands on a to-be-inferred
attribute is low (i.e., all competing brands perform at a near-equal level), the same level may be inferred to exist in brands that lack explicit information about the attribute. We agree that low other-brand variance may serve as a compelling basis for inference, ceteris paribus.

It is less clear how consumers will react when low other-brand variance occurs in the context of other same-brand attributes. Consistent with the preceding experiments, experiment 4 examines the influence of low other-brand variance in the face of competing intuitive relationships. We examine only the low-variance case, because it provides the strongest test of the effects of theory and because our previous experiments have already addressed the high-variance case.

Method

A four-group design that manipulated the presence and nature of the durability-warranty relationship was used. During the study phase, subjects in all conditions viewed descriptions of four fictitious brands of cameras (A–D). These descriptions were identical across conditions except for their treatment of warranty (see Table 4). Durability, which was to be inferred in the test-phase brands, was invariant across the four study brands. That is, all four study brands were given a rating of 11, on a 20-point scale. This represents our operationalization of low other-brand variance.

The four conditions in the present experiment differed from each other in the study phase, test phase, or both, which was dependent on how the intuitively related attribute (warranty) was manipulated. In the high warranty variance condition, warranty varied significantly across brands A–D. Thus, given the lack of variance in durability, the intuitive relationship between warranty and durability was disconfirmed. In the low warranty variance condition, the warranties for brands A–D were as stated. In the no warranty variance condition, warranty information for brands A–D was deleted. In each of these conditions, the warranties for brands E and F were as stated. In the no warranty control, warranty information was deleted from all six brands.

The four conditions in the present experiment differed from each other in the study phase, test phase, or both, which was dependent on how the intuitively related attribute (warranty) was manipulated. In the high warranty variance condition, warranty varied significantly across brands A–D. Thus, given the lack of variance in durability, the intuitive relationship between warranty and durability was disconfirmed. In the low warranty variance condition, the warranties for brands A–D were as stated. In the no warranty variance condition, warranty information for brands A–D was deleted. In each of these conditions, the warranties for brands E and F were as stated. In the no warranty control, warranty information was deleted from all six brands.

<table>
<thead>
<tr>
<th>Warranty A–D condition</th>
<th>Warranty E and F condition</th>
<th>Warranty A–D condition</th>
<th>Warranty E and F condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warranty in study phase</td>
<td>Warranty not used in study phase</td>
<td>Warranty in study phase</td>
<td>Warranty not used in study phase</td>
</tr>
</tbody>
</table>

in the pure control condition, no warranty information was presented for any of the brands. In all conditions except the control, test brand E was assigned a warranty of 12 months and test brand F was assigned a warranty of 24 months. However, in all other respects, brand E was superior to brand F.

A strict interpretation of Ross and Creyer’s (1992) model argues that subjects in all conditions should assign a durability rating of 11 to both test brands, because none of the brands in the study phase deviated from that level. The four experimental conditions offer multiple tests of this prediction. The control condition presents the most straightforward test. Subjects in this condition viewed no warranty information at all. Thus, the durability information gleaned from brands A–D does not compete with theory, and the prediction of the Ross and Creyer (1992) model should be supported. Similarly, in the high warranty variance condition, the model should be supported, because any temptation to infer durability from warranty in the test brands should be eliminated by the obvious lack of correspondence between warranty and durability in the study brands.

If anything, the persuasiveness of other-brand information should increase in this condition because the stimuli show that durability is invariant in the face of an intuitively related predictor.

The no warranty A–D condition provides a more neutral test of the model. In this condition, subjects viewed invariant durability but no warranty during the study phase. Thus, the relationship was neither supported nor disconfirmed. At the time of the test, war-
The proportion of subjects employing each of the three possible inference heuristics is presented in Table 5. Group comparisons reveal that the three heuristics were employed differentially across conditions ($\chi^2 = 7.46, p < .03$). As expected, the other-brand heuristic was prominent in the high warranty variance and control conditions but infrequent in the low warranty variance and no warranty A–D conditions. Similarly, theory served as a basis for inference less often in the control condition than in any of the other conditions.

However, the most revealing findings stem from within-group analyses. The least surprising finding is that subjects in the low warranty variance condition did not rely on the heuristics equally ($\chi^2 = 7.68, p < .03$) but instead were most persuaded by theory. Although the Ross and Creyer (1992) model argues that other-brand heuristics should dominate all conditions, the lack of durability variance in this condition was accompanied by a lack of warranty variance. When inferences were generated for the test brands, theory was still viable and, in fact, was supported by a constant relationship between the critical attributes during the study phase. Thus, while the model failed in this condition, the test of the model may have been inordinately strong.

A fairer test occurs in the no warranty A–D condition. Subjects viewed invariant durability information but no warranty information during the study phase. Thus, the intuitive relationship was neither affirmed nor disconfirmed. During the test phase, warranty information was presented for the first time. Similar to the low warranty variance condition, results for the no warranty A–D condition showed a dominance for the theory heuristic over the other heuristics ($\chi^2 = 12.09, p < .01$). That is, theory dominated despite the (i) invariant durability across brands A–D, (ii) lack of evidence for the theory, and (iii) the presence of a contradictory evaluative consistency cue. The lack of difference between this condition and the preceding one argues that the reliance on theory in the low warranty variance condition was due largely to the effects of theory rather than perceptions of correlation.

Subjects in the high warranty variance condition moderated their reliance on theory as predicted. The surprising result, however, is that theory was no less
influential than other-brand information ($X^2 < 1$). Approximately half of these subjects relied on theory despite (i) invariant durability across brands A–D, (ii) differences in overall performances of the test brands that ran strongly counter to the predictions of theory, and, most important, (iii) empirical disconfirmation of a relationship between warranty and durability. According to the model, dominance of other-brand information should arise in this condition. It should be emphasized that subjects were included in the analysis only if they could recall the durabilities of the study brands.

Finally, the control condition prompted significant reliance on other-brand information and little reliance on theory, as predicted by the model ($X^2 = 4.57$, $p < .05$). The relative influence of these sources of inference is to be expected, given that theory (warranty) was completely absent. More surprising is the level of reliance on evaluative consistency, which was no less than the level of reliance on other-brand information ($X^2 < 1$).

Overall, the results identify theory as the dominant diagnostic cue. In two conditions, theory was a clear winner over other-brand information and evaluative consistency. In a third condition, theory remained an influential force, despite rather obvious empirical disconfirmation of its validity. The only occasion in which theory was not a factor was when it was precluded from being one (i.e., the control condition). The results run contrary to models that argue in favor of mean or discounted mean inferences. In the context of prior research, the predictions of the Ross and Creyer (1992) model are compelling. Much prior research argues in favor of mean inferences, and Ross and Creyer (1992) add the very reasonable assumption that inferences are most likely to revolve around the mean of other brands when variance around that mean is low. However, when allowed to compete, theory and evaluative consistency may be more influential than the mean of other-brand information.

**EXPERIMENT 5**

Given the robust influence of theory in the preceding experiments, we ventured to find potential boundary conditions. In so doing, we were guided by a very pragmatic question concerning the extent to which consumers will be persuaded by a firm’s attempt to use theory to obscure other, less favorable signals pertaining to an inferable attribute. For example, consider the case of Chrysler, which for many years suffered from a quality comparison to Japanese imports. In an attempt to win back consumers, Chrysler offered an unprecedented warranty on its products. The question is whether such theory-based cues can override the effects of history. A priori, it is difficult to predict whether the intuitive relationship between warranty and durability can withstand the effects of contradictory historical performance. History may be persuasive for two reasons. First, it consists of empirical information about the relationship between durability and time. Second, it is possible to construct a causal explanation of the relationship (based, for example, on company strategy or mismanagement). By the inclusion of historical performance as a cue, this experiment differs from the preceding experiments, as well as most other previous research on inference. The failure to consider brand history represents a large gap in our understanding of inference, because it ignores a seemingly compelling basis for inference, that is, brand reputation.

In the present experiment, we pit brand reputation against an intuitive attribute relationship as a basis for inference. Brand reputation was created by providing historical descriptions of the brands over three years. Overall evaluations of the brands were held constant; however, durability differed across but not within brands. Thus, one brand possessed a consistent advantage over the other in terms of durability, but the durability of each brand was relatively constant over time. This design provides an analog of the situation examined by Ross and Creyer (1992). Whereas they examined low variance in a particular attribute across brands, we examine low variance for a particular attribute within brands by including a temporal dimension. We see no reason why low variance in an attribute should be less diagnostic as a cue when it occurs within a brand than when it occurs across brands.

**Method**

The same general procedure was employed as in the past. All subjects were asked to envision a gift-giving scenario and were exposed to the brand information presented in Table 6. This information was said to have been ascertained from an independent magazine source and included the magazine’s own ratings of the durabilities of the brands. As Table 6 indicates, a three-year historical profile of three brands was described. Rank-order and absolute performance of the brands on three attributes—shutter speed, aperture, and program modes—varied over time. Performance of the brands on the fourth attribute, durability, varied across brands and over time in absolute terms, but the same rank order was observed from year to year.

After examining this information, subjects were then provided with information about the 1993 models of the three brands (see Table 6). It was noted that the brands were too new to have been independently evaluated, and therefore only the manufacturer’s specifications were available. This meant that information was available regarding the three objective attributes, but no information about durability was available. Subjects in the control condition were asked to estimate durability on the basis of the three objective attributes. Subjects in the theory condition were given the same information along with the warranties of each brand. The warranties were inversely related to the historical performance of the brands on durability.
INTUITIVE BELIEFS IN INFECTION MAKING

TABLE 6
STIMULI FOR EXPERIMENT 5

<table>
<thead>
<tr>
<th>(Study Phase)</th>
<th>1990</th>
<th>1991</th>
<th>1992</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand X</td>
<td>Brand Y</td>
<td>Brand Z</td>
<td>Brand X</td>
</tr>
<tr>
<td>Maximum Shutter Speed:</td>
<td>5000th</td>
<td>4000th</td>
<td>1000th</td>
</tr>
<tr>
<td>Maximum Aperture:</td>
<td>f1.4</td>
<td>f1.8</td>
<td>f1.2</td>
</tr>
<tr>
<td>Program Modes:</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Durability:</td>
<td>10</td>
<td>13</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(Test Phase)</th>
<th>1990</th>
<th>1991</th>
<th>1992</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand X</td>
<td>Brand Y</td>
<td>Brand Z</td>
<td>Brand X</td>
</tr>
<tr>
<td>4000th second speed</td>
<td>2000th second speed</td>
<td>2000th second speed</td>
<td>f1.4 maximum aperture</td>
</tr>
<tr>
<td>24 month warranty</td>
<td>12 month warranty</td>
<td>12 month warranty</td>
<td></td>
</tr>
<tr>
<td>Durability</td>
<td>Durability</td>
<td>Durability</td>
<td></td>
</tr>
</tbody>
</table>

Note: During the test phase, the presence of the warranty information was manipulated.

Results and Discussion

Brand Z had historically been associated with the highest durability but possessed the shortest 1993 warranty. Thus, subjects were categorized in terms of whether they assigned their highest durability rating to brand Z. Two subjects, one in each condition, did not assign unique ratings to each brand and were discarded from the analysis. Of the remaining subjects, 12 of 15 (80 percent) in the control condition and 10 of 14 (71 percent) in the theory condition assigned their highest durability rating to brand Z ($X^2 < 1$). Postjudgment rationales were consistent with the use of brand history as the basis for selection of brand Z. Although there was a trace of influence of the interattribute relationship, it is clear that the intuitive relationship between warranty and durability did not have the same dominating effect it had in experiments 1–4. In particular, low same-brand variance appeared to compete very successfully against the influence of theory (see experiment 4). This result is comforting from a pragmatic perspective. It appears that warranty cannot be used with impunity by manufacturers who wish to convey high performance on a credence attribute, especially when their reputation suggests otherwise.

Finally, although the results of experiment 5 may be viewed as a potential boundary on the effects of a warranty cue, they do not serve as a compelling boundary on the effects of theory per se. We suspect that brand history was an effective competitor to warranty precisely because subjects were able to construct a “story” about the relationship of history to the durability criterion.

GENERAL DISCUSSION

Research on consumer inference has had a long but checkered history. To this point, the result has been a great deal of uncertainty about the role of inference in consumer decisions. In the present research, we have emphasized the influence of theory as an inference cue. Experiments 1–4 demonstrated the dominance of theory over highly plausible, data-based cues. In particular, experiments 1 and 2 showed that theory can dominate the empirical correlation among attributes as a basis for inference; experiment 4 showed that theory competes successfully against low other-brand variance on an attribute. Experiments 3 and 3A showed how theory can result in spontaneously generated inferences. Experiment 5 showed that a strong theory about an interattribute relationship may not overcome the data if an alternative account of the data can be offered.

Our results concerning the influence of theory are consistent with findings reported in other domains. In addition to the research noted earlier regarding the importance of theory in the perception of causality, recent studies of categorization suggest that intuitive relationships may play a more influential role in classifying objects than does mere feature overlap (see Medin 1989). In fact, when forming categories, people may fail to employ the correlations that exist among attributes unless those correlations can be justified on a conceptual basis (Medin, Wattenmaker, and Hampson 1987). Likewise, in the domain of brand extension it has been shown that evaluation of an extension may depend less on the overall similarity between the original brand and its extension than on inferences consumers draw about the relevance of the brand's characteristics to the benefits desired from the extension category (Broniarczyk and Alba 1994b).

Our research also is in keeping with a popular general model of decision making that argues that decision inputs will influence a decision outcome insofar as they are accessible and diagnostic with respect to the goal of the decision (Feldman and Lynch 1988). The accessibility component of the model has been the recipient of much attention both within and outside the domain of decision making. Many determinants of accessibility have been documented in the memory literature and their effects on decision making are becoming increasingly predictable (see Alba, Hutchinson, and Lynch 1991; Feldman and Lynch 1988). In the context of inference making, Dick et al. (1990) have convincingly demonstrated that the use of an inference rule will vary with its accessibility and the accessibility of the attributes that constitute its arguments.
On the other hand, diagnosticity is relatively unexplored and dominance relations are difficult to predict. Thus, most of our effort was devoted to mapping these relations rather than understanding the role of accessibility or documenting the incidence of spontaneity. Our results, however, are not mute with respect to the latter issue. We do find that high levels of spontaneity can be produced under favorable conditions.

We can only speculate about the low levels of spontaneity reported in previous studies, but we are inclined to believe that spontaneity is a consequence not only of the inference cues but also of method (see Hansen and Zinkhan 1984). Specifically, the judgment task employed in the traditional paradigm may emphasize decision accuracy to a lesser extent than does a choice task, which implies the existence of an optimal decision. Consequently, motivation may be higher in the latter case.

Motivation by itself, however, appears insufficient to induce spontaneity. The present studies, along with other recent research, suggest that the joint criteria of motivation and diagnosticity must be satisfied (see experiment 3A). Neither the presence of diagnostic cues in a judgment task (Simmons and Lynch 1991) nor the presence of correlational cues in a choice task result in a high degree of spontaneity (Dick et al. 1990).

This conclusion notwithstanding, it may now be appropriate to shift attention away from the question of spontaneity. As noted at the outset, inference researchers have diligently probed the situational conditions that facilitate or inhibit the occurrence of spontaneous inference. In contrast, relatively little attention has been devoted to understanding the relative and absolute influence of the multitude of cues that may serve as the basis for inference making. In this sense, we regard the present studies as exploratory. Although experiments 1–4 demonstrate a robust influence of theory, experiment 5 suggests boundary conditions worthy of further research. A natural direction to pursue would be to examine other boundary conditions. Although theory competed effectively against data in the present studies, reliance on one’s intuitive beliefs at the expense of data should be moderated by situational factors such as time pressure, motivational factors such as task importance, and individual differences such as the confidence with which one’s beliefs are held (Kruglanski 1989; Kruglanski and Freund 1983; Maheswaran, Mackie, and Chaiken 1992). A second natural direction to pursue would be to examine an expanded set of inference cues. The present research explored the relative influence of some obvious inference cues but ignored the absolute and relative diagnosticities of many others. For example, even within the domain of belief-based inferences, we examined intuition regarding attribute-to-attribute relations while ignoring category-based relations and similarity-based reasoning (see, e.g., Sujan 1985). Furthermore, recent work suggests that an examination of inference processes should be cognizant of the hierarchical level of cues and the directional nature of inferring the missing attribute (Beike and Sherman 1994). Given the plethora of cues attendant with any product or purchase situation, the need and potential for additional research is virtually unlimited.

[Received June 1993. Revised January 1994. Kent B. Monroe served as editor for this article.]

REFERENCES


Fishbein, Martin and Icek Ajzen (1975), Belief, Attitude, Intention, and Behavior, Reading, MA: Addison-Wesley.


