

Two Ways of Learning Brand Associations

STIJN M. J. VAN OSSELAER
CHRIS JANISZEWSKI*

Four studies show that consumers have not one but two distinct learning processes that allow them to use brand names and other product features to predict consumption benefits. The first learning process is a relatively unfocused process in which all stimulus elements get cross-referenced for later retrieval. This process is backward looking and consistent with human associative memory (HAM) models. The second learning process requires that a benefit be the focus of prediction during learning. It assumes feature-benefit associations change only to the extent that the expected performance of the product does not match the experienced performance of the product. This process is forward looking and consistent with adaptive network models. The importance of this two-process theory is most apparent when a product has multiple features. During HAM learning, each feature-benefit association will develop independently. During adaptive learning, features will compete to predict benefits and, thus, feature-benefit associations will develop interdependently. We find adaptive learning of feature-benefit associations when consumers are motivated to learn to predict a benefit (e.g., because it is perceived to have hedonic relevance) but find HAM learning when consumers attend to an associate of lesser motivational significance.

Associations play an important role in consumers' product evaluations and choices. Brand associations are fundamental to our understanding of inference making (Alba, Hutchinson, and Lynch 1991), categorization (Sujan 1985), product evaluation (Broniarczyk and Alba 1994), persuasion (Greenwald and Leavitt 1984), and brand equity (Keller 1993, 1998). Fundamental to all of these literatures is the assumption that consumers use brand names and product attributes as retrieval cues for information about product performance. In effect, brand names and product attributes are the links to diagnostic information about the product (Feldman and Lynch 1988; Hutchinson and Alba 1991).

Two classes of models have been advanced to explain how associations between brand names and benefits form, change over time, and effect consumer decisions. The first class of models has been championed by Keller (1993, 1998) and can be traced to Anderson and Bower's (1973) Human Associative Memory (HAM) theory (see also Anderson 1983, 1993; Anderson and Lebiere 1998). According to the HAM theory, declarative knowledge is represented as a network of concept nodes connected by links that are strength-

ened each time two events co-occur. Thus, the more a brand name co-occurs with a benefit, either through indirect or direct experience, the stronger the link between the brand name and the benefit. The second class of models, adaptive network models, have been championed by van Osselaer (Janiszewski and van Osselaer 2000; van Osselaer and Alba 2000) and can be traced to the classical conditioning literature (e.g., Gluck and Bower 1988; Rescorla and Wagner 1972; Rumelhart, McClelland, and the PDP Research Group 1986). According to these adaptive network models, association strengths update and evolve as cues interact, and often compete, to predict outcomes. Thus, whereas HAM models hold that cues are learned independently, adaptive network models hold that cues interact. That is, the strength of the association between a brand name and a benefit depends on how uniquely a brand name can predict the benefit.

To date, there has been a limited appreciation of the relevance of these two classes of models for understanding brand associations. Consumer learning research has usually treated products as single, aggregate objects instead of conglomerates of features such as family brand names, subbrand names, and attributes. When products are perceived in the aggregate, cue interaction is not an issue and both classes of models make largely the same predictions. Even the few consumer learning studies that have presented products as multiple-cue stimuli (e.g., Hutchinson and Alba 1991; Meyer 1987) were not designed to separate the two classes of models. Recently, however, Janiszewski and van Osselaer (2000) designed a series of five experiments that explicitly contrasted the two classes of models in a brand association

*Stijn M. J. van Osselaer is assistant professor of marketing, University of Chicago, 1101 East 58th Street, Chicago, IL 60637 (stijn.vanosselaer@gsb.uchicago.edu). Chris Janiszewski is the Jack Faricy Associate Professor of Marketing, University of Florida, Gainesville, FL 32611-7155 (chrisj@dale.cba.ufl.edu). This work was supported by the James M. Kilts Center for Marketing at the University of Chicago Graduate School of Business. The authors wish to thank Alan Cooke, Rich Lutz, and Ann McGill for their helpful comments. Both authors contributed equally to this article.

context and found strong cue interaction effects. Van Osselaer and Alba (2000) also found a persistent cue interaction effect (blocking) and showed that this effect could not be explained by a backward-looking causal reasoning process. When this consumer evidence is combined with the growing number of demonstrations of cue interaction in the category, causal and multiple-cue probability learning literatures (e.g., Chapman and Robbins 1990; Dickinson, Shanks, and Evenden 1984; Gluck and Bower 1988; Kruschke 1996; Kruschke and Johansen 1999; Lopez et al. 1998; Shanks 1991), one is tempted to conclude that cue interaction might be universal in consumer learning of product associations. This would imply either that HAM models need to be extended to explain cue interaction phenomena or that all product associations are learned through an adaptive process that is best described by adaptive network models. The latter would imply that HAM theory is no longer a viable explanation of how product associations are formed, updated, and, ultimately, lead to evaluations and choices.

The goal of this article is to show that HAM models and adaptive learning models describe two unique approaches to learning the brand associations that are subsequently used to make predictions about consumption benefits. We provide insight into the qualitative differences between the processes responsible for each type of learning and predict when each type of learning is most likely to drive consumers' product evaluations. Study 1 shows that consumers adaptively learn associations between brand names and product benefits when a target benefit is the focus of prediction during learning but that they engage in HAM learning of a secondary benefit that is not the focus of prediction. Study 2 rules out the possibility that the learning system is constrained to adaptive learning of a single benefit by showing that consumers can adaptively learn about two benefits, provided both benefits are the focus of prediction during learning. Study 3 shows that merely paying attention to an outcome is insufficient to produce adaptive learning. Rather, an outcome has to be significant enough for the learning system to be motivated to focus on learning how to predict that outcome. Study 4 shows that people can be discouraged from relying on adaptively learned associations and instead rely on HAM learned associations when they are asked to retrospect about their product experiences. Together, the studies provide support for a two-process system that has HAM learning as its foundation but engages in adaptive learning when attention is directed at significant outcomes.

MODELING ASSOCIATIONS

Models of associations in memory attempt to explain two cognitive events. First, why do cues activate some concepts but not others? Second, why does the likelihood that a cue will activate a concept change over time? Insight into these issues relates directly to our understanding of how consumers retrieve brand names, product attributes, and sources of marketing communications from memory (e.g., Alba et al. 1991; Nedungadi 1990; Pham and Johar 1997). Associative models also inform us about why extending a brand name

to lower quality products hurts brand equity (Loken and John 1993), why brands can extend to some categories but not others (Boush and Loken 1991; Broniarczyk and Alba 1994; Park, Milberg, and Lawson 1991), why brand extensions can hurt brand beliefs but not flagship product beliefs (John, Loken, and Joiner 1998), and why brand alliances can create expectations of a superior product (Park, Jun, and Shocker 1996) and can benefit a weaker partner (Simonin and Ruth 1998). Thus, understanding how associations between brand names and benefits form and change over time is a fundamental research issue.

Two general classes of models have been advanced to explain how the associations that consumers use to predict benefits develop over time. The first class of models have evolved from the HAM theory of human cognition (Anderson and Bower 1973). The second class of models have evolved from the Rescorla-Wagner model of conditioned learning (Rescorla and Wagner 1972). Interestingly, the models make both convergent and divergent assumptions about the learning processes responsible for changes in the strength of associations between cues and outcomes. We provide a general description of the two classes of models and their similarities and differences below, as well as verbal descriptions of two representatives of each class of models, chosen because they are the most current, prototypical, and empirically supported in their respective literatures.

The HAM Models

Two types of models can be distinguished within the HAM tradition, connectionist models and Bayesian models (see App. A for detailed review). These models have similarities and differences. With respect to similarities, both types of models represent declarative knowledge as a network of concept nodes connected by associative links. In these models, the association strength between a cue and an outcome, as well as the activation of the outcome on presentation of the cue, depend critically on the frequency of co-occurrence between the cue and the outcome. These models assume that the learning of associations between one cue and an outcome is not dependent on the presence of other cues and their associations with the same outcome. That is, learning is cue independent.

The primary difference between the connectionist and Bayesian models is in the way that they represent learning. Connectionist models employ a simple Hebbian or "coincident activation" learning rule (e.g., McClelland 2000; Rumelhart et al. 1986). One example of these models is the simple Direct Association (DA) model described by Janiszewski and van Osselaer (2000), in which each cue is connected to each outcome, each connection is strengthened whenever a cue and an outcome co-occur, and incoming activation to an outcome node is combined additively. Bayesian models take the form of statistical Bayesian inference. The prime example of these models is Anderson's Adaptive Control of Thought-Revised (ACT-R) model (Anderson 1993; Anderson and Lebiere 1998). According to the ACT-R model, the activation of an outcome node is a function

of (1) that node's base level of activation and (2) incoming activation to the outcome node from cueing nodes. Base-level activation is a function of the frequency and recency of the outcome's occurrence. Incoming activation is an additive sum of association strengths from cues to the outcome, weighted by the activation of the cues. Association strengths, in turn, are dependent on a number of factors, including preexperimental priors, whether the outcome is more likely to occur in a cue's presence than in that cue's absence, and, importantly, frequency of cue-outcome co-occurrence. It is also important to note that, according to the ACT-R model, learning of cue-outcome associations is not affected directly by the presence of other cues and their associations with the same outcome. Thus, like the other HAM models, ACT-R association strengths depend on the frequency with which a cue and an outcome occur together, regardless of the associations that other, also-present cues have formed with the same outcome.

Adaptive Network Models

Learning in the classical conditioning tradition, and in the category, causal, and multiple-cue probability learning work based in this tradition, is often described using adaptive network models (e.g., Gluck and Bower 1988; Kruschke and Johansen 1999; Pearce 1994). Like HAM models, these models represent knowledge as a network of nodes connected by associative links. Moreover, like HAM models, activation of a noninput node is a sum of incoming activations. Finally, like the Hebbian HAM models, the models are connectionist in nature.

Two examples of adaptive network models that have found considerable empirical support are the Least Mean Squares (LMS) model (Gluck and Bower 1988) and Pearce's (1994) configural model (see App. A for detailed review). In the basic LMS model (e.g., Gluck and Bower 1988; Janiszewski and van Osselaer 2000), a single layer of input nodes is connected to a single layer of output nodes, and the connections are updated using the simple Delta error reduction learning rule. Pearce's (1994) configural model adds two hidden layers of nodes between input and output nodes. One extra layer allows the model to explain a reduction of attention to each cue when multiple cues are present. The other extra layer represents configurations of stimulus elements, leading the model to predict sensitivity to specific combinations of stimulus elements.

The distinguishing feature of both types of adaptive network models is the error-driven learning rule. That is, updating of associations only takes place to the extent that the learning system is not already correctly predicting an outcome. Thus, increasing the frequency of cue-outcome co-occurrence does not necessarily increase the strength of the association between a cue and an outcome. In addition, because predicted outcomes and, hence, errors are influenced by all present cues (because of the fact that output activation is a sum of incoming activations), updating one cue's association with an outcome depends on the strength of the associations of other, copresent cues with the same outcome.

Thus, the cue additivity and error reduction properties together create a cue-interdependence or cue-interaction property—the associations between multiple cues and an outcome are learned interdependently.

Two important examples of cue-interactive learning are the blocking and unblocking phenomena. If a cue has a strong association with an outcome (i.e., predicts it will occur), simultaneous presentations of an additional cue will not lead that cue to develop a strong association with the outcome, regardless of its frequency of co-occurrence with the outcome. The predictive ability of the first cue blocks learning about the second cue. In contrast, if a cue has a negative association with the outcome (i.e., predicts it will not occur), but the presentation of an additional cue leads to the outcome, then the association between the additional cue and the outcome will have to be strongly positive to make the sum of the activation from the combined cues positive. The additional cue is unblocked. Together, the blocking and unblocking phenomena lead to situations in which an unblocked cue that co-occurs with an outcome less frequently than a blocked cue still ends up with a stronger outcome association.

TWO ASSOCIATIVE LEARNING PROCESSES

Although recent evidence is consistent with the view that adaptive network models can account for all associative learning, it is also possible that there are two distinct types of associative learning processes that can both drive evaluations and choices. The first is a process characterized by cue independence that is best described by models in the HAM tradition. The second is a process characterized by cue interaction that is best described by adaptive network models. The possibility that product associations are not always learned through an adaptive learning process is consistent with the existence of a small number of examples in the literature showing a lack of cue interaction. However, the ability of these examples to disprove the universality of adaptive learning has been disputed, for example, because the lack of interaction between outcomes was wrongly framed as a lack of interaction between cues (Waldmann and Holyoak 1992; see also Shanks et al. 1996; Shanks and Lopez 1996; Van Hamme, Kao, and Wasserman 1993), because chunking made multiple-cue stimuli into single-cue stimuli (Williams, Sagness, and McPhee 1994), or just because evidence of the absence of cue interaction was based on null-effect findings (e.g., Baeyens et al. 1996; Baeyens et al. 1998). Thus, whether or not people have two associative learning processes that can be used to make outcome predictions is a timely issue.

Adaptive Learning

Our goal is not only to propose that two processes are involved in associative learning but to provide some insight into the qualitative nature of each process and situations in which each process is likely to be dominant. Closer ex-

amination of the learning rule that is central to the adaptive network models (see App. A) yields insights about the type of process they describe. The structure of their learning equation implies a specific type of process—one that is essentially forward looking and adaptive. First, a prediction is made about an outcome. Then, feedback is received about the outcome in the form of the actually experienced level of the outcome. Next, association strengths are updated to improve prediction on the next occasion. This prediction-feedback-update process continues until prediction is perfect. The process is directional and geared toward optimizing a prediction on the next occasion. Associations go from cues to outcomes so that activation can flow forward for the purpose of prediction. Feedback may flow backward, but activation does not. This implies that judgments about an outcome on the basis of information about a cue do not necessarily yield the same answers as judgments about a cue on the basis of an outcome (see Farquhar and Herr [1993] and Price and Yates [1995] for examples of such associative asymmetry).

The forward-looking adaptivity of the system also leads to association strengths that do not reflect each experience with a cue and an outcome equally. The extent to which a co-occurrence between a cue and an outcome is reflected in the strength of their association depends heavily on when that co-occurrence takes place and on the presence of other cues. This is reflected in findings of strong order effects. For example, Janiszewski and van Osselaer (2000) asked subjects to taste a series of eight samples of baked goods and to rate their taste quality. The superior tasting samples were the third, fourth, seventh, and eighth samples and were labeled “Treats with Nu-Oil,” “Goodies,” “Treats,” and “Goodies with Oilean.” They found that the ingredient brand Nu-Oil became more associated with the benefit of good taste than the ingredient brand Oilean, even though each subbrand described exactly one superior tasting product. According to the adaptive network models, subjects learned to predict that a product would taste superior based on their initial positive experience with the Goodies product; hence, the Oilean brand name was blocked from acquiring associative strength on the subsequent experience with the Goodies with Oilean product. In contrast, the initial experience with the Treats with NuOil product allowed subjects to associate both brand names with superior taste because both were novel and did not already predict superior taste.

Van Osselaer and Alba (2000, experiment 2B) provide additional evidence that cue interaction results from a forward-looking learning system. In their experiment, they created a situation in which the retrospective frequencies of an outcome given two cues were the same for both cues. Thus, a backward-looking process should lead to both cues being learned equally well, and no cue interaction should occur. However, they also created the situation such that a forward-looking process would lead to cue interaction. They found a strong cue interaction effect, indicating that consumer learning of brand associations is forward looking.

Finally, it is important to recognize that adaptive learning

is a selective process. Outcomes have a different status than cues. Cues are used to predict outcomes, not the other way around. Thus, some distinction has to be made between elements that are used as predictors and elements that are used as the outcome to be predicted. It is very well possible that such a distinction is made automatically, but it does suggest a (automatic or nonautomatic) motivation to focus on predicting some elements and not others. We hypothesize that for adaptive learning (and, hence, cue interaction) to occur, consumers need to focus on a stimulus element as outcomes to be predicted. We also expect that predictive focus requires motivation, which should depend heavily on the perceived hedonic relevance of the outcome. For example, when learning about a product category, consumers should be more likely to focus on predicting characteristics of the consumption experience that are perceived to have rewarding or punishing implications (i.e., the taste experience provided by a food item) than on characteristics of the product that are less directly linked to the quality of the consumption experience (i.e., ingredients of a food item).

HAM Learning

The HAM models seem to reflect a qualitatively different process than the adaptive network models. In HAM models, there are no predictions made that are tested and compared with feedback information. Instead, learning reflects the basic Hebbian learning principle of “what fires together, wires together.” In addition, there is no real difference between the different stimulus elements. For example, it is essentially arbitrary which stimulus element is designated as cue or outcome in the DA model, as they have exactly the same function in the model. These characteristics of HAM-based models suggest that the process underlying HAM models might be less focused than adaptive learning.

We believe there are many situations in which consumers are not focusing on predicting a particular element of a stimulus. Consumers may lack the motivation to predict outcomes, may have difficulty identifying a relevant outcome, or may simply be in a situation where a relevant outcome does not exist. In other situations, consumers may be trying to predict one outcome, but not others, about which information is also available. In such situations, it may be better simply to store elements of the experience. In this storage process, all elements are cross-referenced by establishing simple associations that are strengthened each time two elements appear together. When consumers later are in a situation in which they need to predict one element based on the presence of a second element, they will retrospectively try to recall what level of the first element was paired with the present level of the second element. Thus, we propose that there are many situations in which consumers use a backward-looking, cue-independent process to make predictions about product performance and that this process is best described by HAM learning models.

There is some evidence for the existence of associations that (1) are not subject to cue interaction and (2) co-exist with other associations that are subject to cue interaction.

Van Osselaer and Alba (2000) asked consumers to learn how to predict the quality of a white-water raft using 16 product profiles. They found that when consumers learned to predict a desired benefit based on one predictive cue in the first four profiles, they failed to learn the predictive value of an additional predictive cue that was introduced later. This is a typical cue interaction phenomenon predicted by adaptive learning models. However, when van Osselaer and Alba (2000, experiment 2A) asked subjects to recollect and report the pattern of co-occurrence between the additional cue and the benefit prior to making quality predictions, the additional attribute had more effect on quality predictions. Thus, the cue interaction effect was attenuated. In hindsight, this finding might be interpreted as evidence for the existence of a second associative learning process. It is possible that the retrospective retrieval instructions encouraged people to use associations formed as a consequence of HAM learning and to ignore associations formed as a result of adaptive learning. However, van Osselaer and Alba (2000) mention that the attenuation result could result from subjects' desire to be consistent in their public judgments of co-occurrence and quality.

In addition to van Osselaer and Alba's (2000) results, which suggest that there are cue-independent associations that can influence product judgments, there is evidence suggesting that the person's focus of learning moderates whether adaptive or HAM learning drives their judgments. Hutchinson and Alba (1991) find that people are more likely to put higher weight on a perfectly predictive cue, and less weight on other cues, when learning is intentional rather than incidental. It is possible that intentional learning instructions made Hutchinson and Alba's (1991) criterion outcome the focus of prediction, making learning about this outcome more adaptive and more directed toward optimal prediction. This may have led subjects to place more weight on the critical cue and less on the other cues. In contrast, under incidental learning instructions, the main outcome may not have been the focus of prediction. In that case, consumers are less geared toward optimal prediction, leading to weights that are less dependent on a cue's critical contribution to optimal prediction. However, we note that Hutchinson and Alba's (1991) result can be explained in a number of ways that are consistent with both types of learning. For example, both classes of models would predict that the association strengths of perfectly predictive cues are more similar to those of imperfectly predictive cues if incidental learning is slower than intentional learning.

In sum, we believe that consumers' predictive product judgments are also influenced by an associative learning process that is qualitatively different from adaptive learning. This process is backward looking—instead of the activation of cues leading directly to judgments or predictions of outcomes, judgments made using the HAM system rely on associative retrieval of information about how often stimulus elements have occurred together in the past. In addition, during learning, the process does not require consumers to focus on predicting specific stimulus elements. Finally, be-

cause learning in this system is based on basic co-occurrence instead of a weighted additive prediction rule and error minimization, we believe that this learning should not be characterized by cue interaction. Thus, we expect that consumers' product judgments will show cue interaction and will be consistent with adaptive network models for outcomes that are the focus of prediction during learning. We expect that product judgments will show cue independence and will be consistent with HAM models for outcomes that are not the focus of prediction during learning. Study 1 examines this hypothesis in a scenario with two outcomes by instructing subjects to focus on predicting one of the outcomes.

STUDY 1

If adaptive learning is contingent on an outcome being the focus of prediction during learning, then it should be possible to encourage consumers to engage in adaptive learning or HAM learning about relationships between brand names and consumption outcomes by varying the focus of prediction during learning. Subjects were asked to taste and rate a series of six cake samples that varied both in the intensity of their chocolate flavor and the degree of their moistness. In addition to receiving information about flavor and moistness, subjects also received information about each product's family brand name and, in some cases, an ingredient brand name. The pairing between the brand name(s) and the outcomes (flavor and moistness) was manipulated to allow us to draw inferences about the nature of the associative learning process between ingredient brand names and each of the two outcomes. One-half of the subjects were encouraged to focus on learning to predict the flavor benefit by having them rate the chocolate flavor of the sample, whereas the other half of the subjects were encouraged to focus on learning to predict the moistness benefit by having them rate the moistness of the sample. Subsequently, subjects were asked to predict the flavor and the moistness of a muffin that had ingredient brand name I_1 or I_2 .

The key dependent measure was the association between the ingredient brand names and chocolate flavor. It was expected that the flavor-focused subjects should engage in adaptive learning of the flavor association as evidenced by cue interaction effects in the associations between the ingredient brand names and chocolate flavor. In contrast, the moistness-focused subjects should engage in HAM learning of the flavor association, as evidenced by cue independence of the associations between the ingredient brand names and chocolate flavor.

Design and Predictions

The study was a two-cell, within-subject design (ingredient brand name that uniquely predicts the target benefit), with a between-subject manipulation of the focus of prediction during learning (target benefit, distracter benefit) and an ingredient brand name counterbalancing factor (see App. B, Tables B1, B2, and B3, for design summaries of this and

the following studies). Subjects were exposed to an L, F₁, F₁I₁, F₂, F₂I₂, F₂I₂ sequence of product experiences in which L represented a low-quality baseline brand, F₁ and F₂ represented family brand names (e.g., Delight, Buon Chocolate), and I₁ and I₂ represented ingredient brand names (e.g., Baker's Blend Syrup, Silk'n Morsels). The cakes differed in their moistness and chocolate flavor. The L and F₁ cakes had a mild chocolate taste. The other cakes had a strong chocolate taste. The L and F₂ cakes tasted dry, whereas the other cakes tasted moist.

According to the HAM models, I₂ should have a stronger association with strong chocolate flavor than I₁, because I₂ was experienced together with strong chocolate flavor twice whereas I₁ had only been paired with strong chocolate flavor once (see Fig. 1 and App. C). Adaptive learning models predict the opposite result. Cake I₁ should have a stronger association with strong chocolate flavor than I₂ because of an unblocking effect (Rescorla and Wagner 1972; Shanks 1991) for I₁ and a blocking effect (Kamin 1969) for I₂. According to the adaptive learning models, the I₁ cue (e.g., Baker's Blend Syrup) will form a strong association with strong chocolate flavor whereas the I₂ cue (e.g., Silk 'n Morsels) will not, despite the fact that both ingredient brand names are presented with strong chocolate flavor (see Fig. 1 and App. C). This is the case because the ingredient brand name uniquely identifies strong chocolate flavor in the first product line but is a redundant predictor for strong chocolate flavor in the second product line (the strong chocolate flavor is already predicted by the family brand name F₂).

The two classes of models do not make competing predictions about the association strength between the two in-

redient brand names and the moistness benefit. The HAM models predict that I₂ will be perceived as moister than I₁ because I₂ is presented with moist cake twice and I₁ is presented with moist cake once. The adaptive network models predict that I₂ will be perceived as moister than I₁ because I₂ uniquely identifies moist taste whereas I₁ is a redundant predictor. Nonetheless, the moistness manipulation was included because subjects in the distraction condition needed to attend to a consumption outcome with variance.

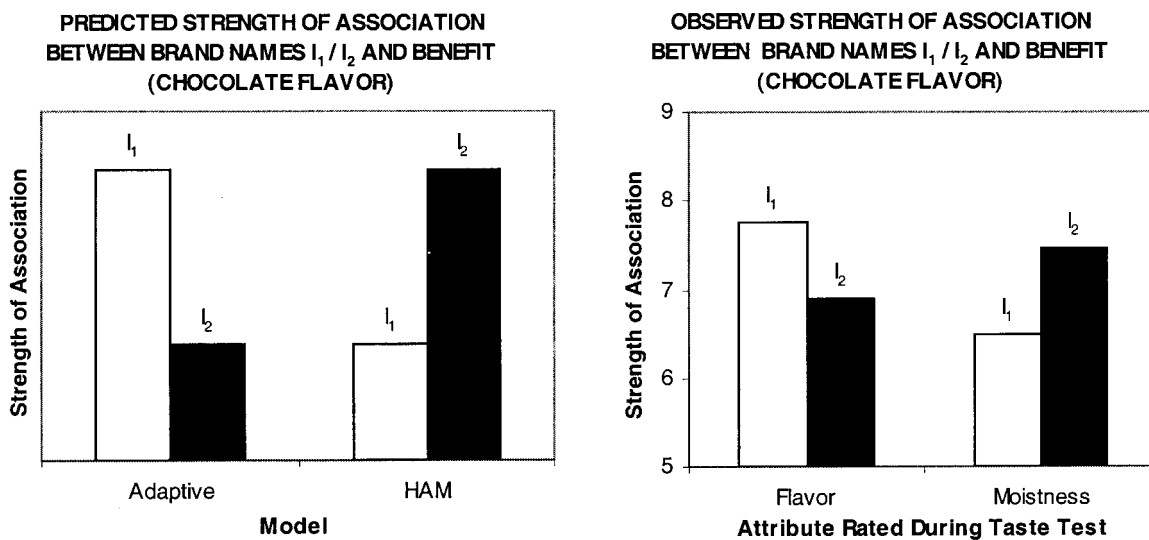
Procedure and Stimulus Materials

Subjects participated in a taste test experiment in which they were asked to rate a series of six cake samples. Subjects were given a questionnaire, the first page consisting of an explanation of the task. Turning to page 2, subjects were asked to taste and rate the first sample on one of two 10-point scales, with end points labeled dry/extra-moist and mild chocolate flavor/strong chocolate flavor. They first tasted the L cake sample labeled Treats. The questionnaire informed them that the Treats cake was the lowest quality sample that they would taste and that it should be rated a "1" on the scale. Subjects were told that the rating of the five remaining cake samples should be made relative to this baseline.

The subjects then tasted the remaining five cake samples and were asked to rate the intensity of the chocolate flavor or the moistness of each sample. Each of the five remaining cake samples had been labeled with brand/branded ingredient information. Each cake sample was also identified with brand/branded ingredient information in the questionnaire.

FIGURE 1

PREDICTIONS AND RESULTS OF STUDY 1



Subjects were asked to check the product label prior to tasting the cake sample to confirm that they were tasting the correct sample. This instruction increased the salience of the brand and branded ingredient information. Subjects ate an unsalted soda cracker and had a drink of water in between each of the five product taste tests. Subjects always rated the current sample prior to tasting the next sample.

The mild, dry baseline I_1 sample was created by taking a store brand chocolate cake, replacing one-half of the mix with white cake mix, and then air-drying the baked cake for 24 hours. The mild, moist F_1 sample was created by taking a chocolate cake mix (e.g., Betty Crocker Super Moist Fudge Chocolate) and replacing one-half of the mix with white cake mix (e.g., Betty Crocker Super Moist White). The flavorful moist F_1I_1 sample was the unaltered chocolate cake mix (e.g., Betty Crocker Super Moist Fudge Chocolate). The flavorful, dry F_2 sample was a base chocolate cake mix (e.g., Betty Crocker Supreme Chocolate), with an extra cup of chopped chocolate chips that was then air dried for 24 hours. Pretesting showed that dry samples were perceived to have less chocolate flavor than moist samples, so chocolate chips were added to compensate for the dryness. These chips were not noticeable in the baked product. The flavorful, moist F_2I_2 sample was the same base chocolate cake mix as the F_2 cake mix (e.g., Betty Crocker Supreme Chocolate) but without the chocolate chips or drying. The second flavorful, moist F_2I_2 sample was a fudge cake (e.g., Betty Crocker Super Moist with Creamy Swirls of Fudge Chocolate). The second F_2I_2 sample was not the same sample as the first, because it would have been odd to our subjects to taste the same thing twice. However, any differences between the two samples beyond their chocolate flavor and moistness do not affect the predictions for either class of learning models. The Baker's Blend Syrup and Silk'n Morsels ingredient brand names assigned to I_1 and I_2 were counterbalanced.

After tasting all of the samples, the sample wrappers and labels were collected for disposal. Subjects were then asked to turn the page and report their expectations about the chocolate flavor and the moistness of a muffin made with each ingredient brand (I_1 and I_2). Subjects used 10-point scales to report how likely it would be that a Muffin Man muffin made with Baker's Blend Syrup (F_3I_1) would have a strong chocolate flavor, how likely it would be that a Muffin Man muffin made with Silk'n Morsels (F_3I_2) would have a strong chocolate flavor, how likely it would be that a Muffin Man muffin made with Baker's Blend Syrup (F_3I_1) would taste moist, and how likely it would be that a Muffin Man muffin made with Silk'n Morsels (F_3I_2) would taste moist. The F_3I_1 and the F_3I_2 chocolate flavor dependent measures were counterbalanced but always were the first two of the four judgments.

Results

Fifty-nine subjects from an introductory marketing class were awarded extra credit to participate in the study. The expectation about the intensity of the chocolate flavor of the ingredient branded muffin was used as an indicator of the

association strength between each ingredient brand name and flavor. There was no influence of the ingredient brand name counterbalancing factor or the dependent measure order counterbalancing factor on the flavor ratings (all p 's > .10).

Manipulation Check. Subjects who rated the flavor of the samples perceived the F_1 sample to have less chocolate flavor than the F_1I_1 sample ($F(1, 28) = 201.75, p < .05$) and the F_2, F_2I_2, F_2I_2 samples to have equivalent chocolate flavor ($F(2, 26) = 1.23, p > .10$). Subjects who rated the moistness of the samples perceived the F_1 sample to be as moist as the F_1I_1 sample ($F(1, 28) = 0.17, p > .10$) and the F_2 sample to be drier than the F_2I_2 and F_2I_2 samples ($F(2, 27) = 31.98, p < .05$). Thus, the manipulation of the two taste dimensions was successful.

Associations with Ingredient Brand Names. Figure 1 shows the degree of association between I_1 and I_2 with chocolate flavor by condition. The interaction of the focus of prediction during the learning phase (i.e., focus on flavor vs. focus on moistness) and the amount of chocolate flavor associated with the ingredient brand names, I_1 and I_2 , was significant ($F(1, 26) = 8.06, p < .05$). When subjects were asked to rate the chocolate flavor intensity of the samples during the learning phase, I_1 was perceived as a better predictor of the chocolate flavor than I_2 ($M_{F_3I_1} = 7.76, M_{F_3I_2} = 6.90, F(1, 27) = 4.06, p = .05$), a finding consistent with the predictions of the adaptive network models. When subjects were asked to rate the moistness of the samples during the taste test, I_2 was perceived as a better predictor of chocolate flavor than I_1 ($M_{F_3I_1} = 6.50, M_{F_3I_2} = 7.47, F(1, 28) = 4.11, p = .05$), a finding consistent with the HAM models. Subjects also perceived a muffin made with I_2 ($M_{F_3I_2} = 7.25$) would be moister than a muffin made with I_1 ($M_{F_3I_1} = 6.30, F(1, 55) = 30.98, p = .05$), a result that is consistent with adaptive network and HAM models.

Discussion

In study 1, people were encouraged to focus on predicting either the flavor or the moistness of cakes. As expected, a predictive focus on the chocolate flavor of the samples led to adaptive learning of associations between the brand names and the flavor of the products, whereas a predictive focus on the moistness of the samples led to HAM learning for the associations between brand names and flavor. Thus, study 1 shows that consumer learning of brand-benefit associations and subsequent product judgments are not necessarily subject to cue interaction. This result suggests that consumers have two different ways of learning product associations that can influence product judgments, one consistent with HAM models of associative learning and memory and one consistent with adaptive network models.

The results of study 1 suggest that the focus of prediction is a critical determinant of which type of associations drive judgments. Yet, it is possible that people focused on both

benefit outcomes but could not learn adaptively about more than one outcome at a time. This alternative hypothesis is consistent with the surprising fact that we are not aware of any published data showing adaptive learning with stimuli that consist of both multiple cues and multiple outcomes. Thus, the finding that adaptive learning was limited to the benefit that was the focus of attention may have been a constraint of the learning system, not a consequence of the predictive focus only. If being the focus of prediction is the critical factor that drives adaptive learning, then it should be possible to find evidence for adaptive learning if consumers focus on two outcomes during learning.

STUDY 2

Study 2 was designed to test the hypothesis that adaptive learning can occur for two outcomes when both are the focus of prediction during learning. To the extent consumers can learn adaptively about more than one outcome, the hypothesis that adaptive learning is not limited to a single outcome, but critically depends on predictive focus, would be further supported.

Design and Predictions

The study was a 2 (ingredient brand name that uniquely predicts target benefit one) by 2 (ingredient brand name that uniquely predicts target benefit two) within-subject design with a between-subject manipulation of the ingredient brand/consumption outcome combination receiving the greater number of pairings (see App. B). The design also included an ingredient brand name counterbalancing factor and a dependent measure order counterbalancing factor.

The same cake samples and dependent measures were used as in study 1. The learning sequence in condition 1 was identical to the learning sequence experienced in study 1. Thus, in condition 1, subjects experienced a mild chocolate-flavored piece of cake F_1 , followed by strong chocolate-flavored cakes F_1I_1 , F_2 , F_2I_2 , and F_2I_2 . In this condition, the HAM models predict that I_2 should be more strongly associated with the chocolate flavor, whereas the adaptive network models predicts a stronger association of I_1 with the chocolate flavor (see Fig. 2). The same series of samples also varied with respect to moistness. The F_1 and F_1I_1 samples were moist, followed by a dry piece of cake F_2 and moist pieces of cake F_2I_2 and F_2I_2 . Both classes of models predict that I_2 should become more strongly associated with moistness (see Fig. 2).

To show that people can engage in the adaptive learning of two benefits, we needed a second condition, because any one condition can only differentiate between two types of learning for a single benefit. In condition 2, subjects experienced a mild chocolate-flavored piece of cake F_1 and then strong chocolate-flavored cakes F_1I_1 , F_1I_1 , F_2 , and F_2I_2 . In this learning scenario, both models predict that I_1 becomes more associated with the chocolate flavor (see Fig. 2). The same pieces of cake also varied in terms of moistness. The F_1 , F_1I_1 , and F_1I_1 samples were moist pieces of cake, fol-

lowed by a dry piece of cake F_2 and a moist piece of cake F_2I_2 . The HAM models predicts that I_1 will be more strongly associated with moistness, whereas the adaptive network models predict that I_2 will be more strongly associated with moistness. All subjects were asked to rate both the chocolate flavor and the moistness of the samples during the taste test, encouraging subjects to predict both benefits during learning.

Results

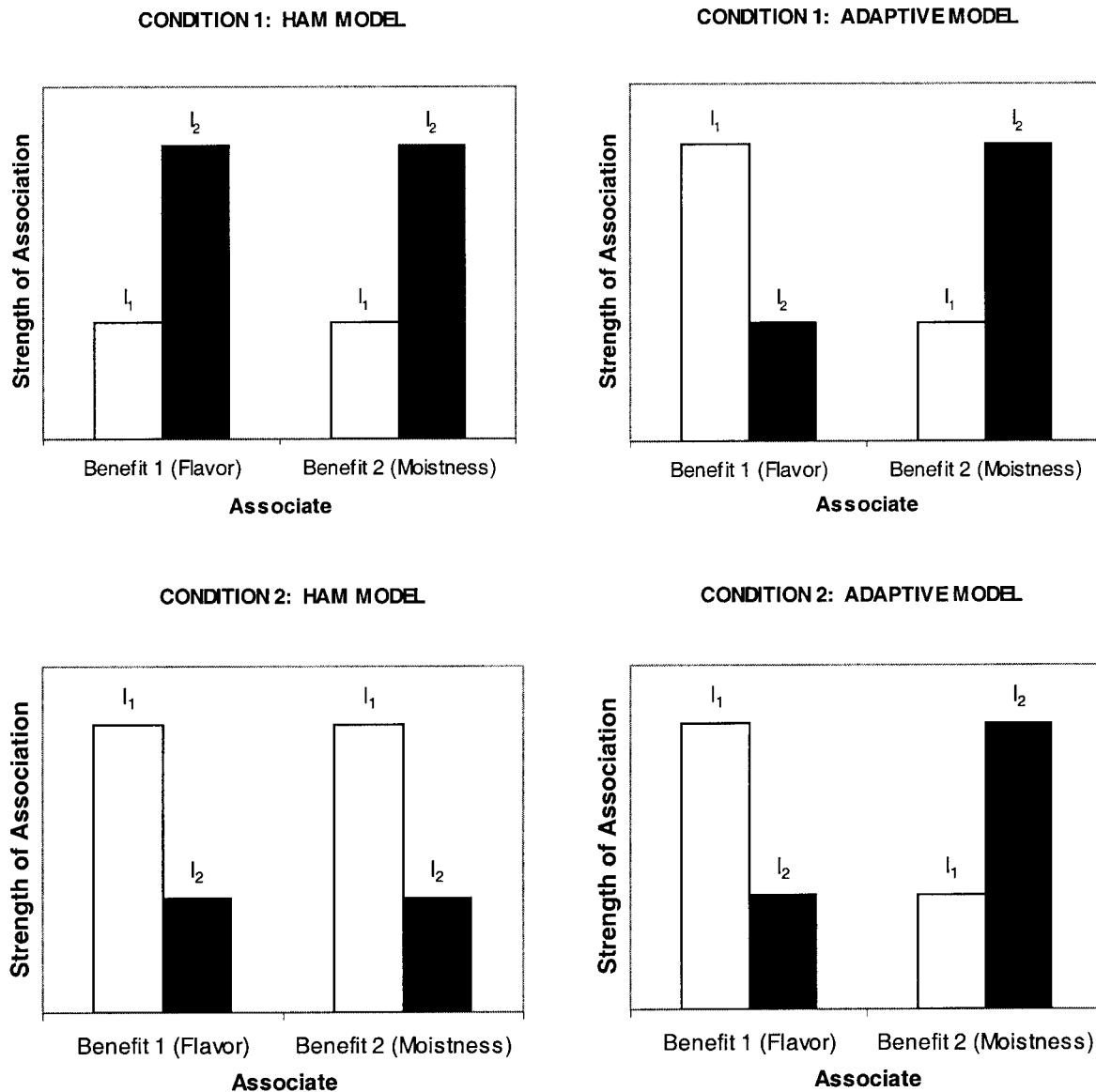
Fifty-eight subjects from an introductory marketing class were awarded extra credit to participate in the study. The expectation about the intensity of the chocolate flavor of the ingredient-branded muffin was used as an indicator of the association strength between each ingredient brand name and flavor. The expectation about the moistness of the ingredient-branded muffin was used as an indicator of the association strength between each ingredient brand name and moistness. There was no influence of the sample rating question order counterbalancing factor for either dependent measure in either condition (all p 's > .10). There was no influence of the ingredient brand name counterbalancing factor for either dependent measure in either condition (all p 's > .10).

Manipulation Check. In condition 1, subjects perceived the F_1 sample to have less chocolate flavor than the F_1I_1 sample ($F(1, 27) = 76.54, p < .05$) and perceived the F_2 , F_2I_2 , and F_2I_2 samples to have equivalent chocolate flavor ($F(2, 25) = 1.59, p > .10$). In condition 2, subjects perceived the F_1 sample to have less chocolate flavor than the F_1I_1 and F_1I_1 samples ($F(1, 29) = 16.79, p < .05$) and perceived the F_2 and F_2I_2 samples to have equivalent chocolate flavor ($F(1, 28) = 0.68, p > .10$). In condition 1, subjects perceived the F_1 and F_1I_1 samples to have equivalent moistness ($F(1, 26) = 0.54, p > .10$) and perceived the F_2 sample to be drier than the F_2I_2 and F_2I_2 samples ($F(1, 27) = 13.80, p < .05$). In condition 2, subjects perceived the F_1 to be moister than both F_1I_1 samples ($F(2, 27) = 6.21, p < .05$) and perceived the F_2 sample to be drier than the F_2I_2 sample ($F(1, 29) = 58.89, p < .05$). The finding that F_1 was moister than both F_1I_1 samples in condition 2 was unexpected, but the higher moistness for the F_1 sample did not affect the directional predictions of either class of learning models. For the cue-independent HAM models, the moistness of a sample that does not include an ingredient brand does not affect that ingredient brand's associations. For adaptive network models, the fact that the F_1 sample is moister than the F_1I_1 sample makes I_1 even less of an indicator of high moistness than if both samples were equally moist, which is perfectly consistent with the prediction that I_1 has a weaker association with moistness than I_2 .

Associations with Ingredient Brand Names. Figure 3 shows the degree of association of I_1 and I_2 with chocolate flavor and moistness by condition. In condition 1, I_1 was perceived as a better predictor of chocolate flavor than I_2

FIGURE 2

PREDICTED STRENGTH OF ASSOCIATION BETWEEN INGREDIENT BRAND NAMES AND BENEFITS IN STUDY 2



($M_{F_{3I_1}} = 7.68$, $M_{F_{3I_2}} = 6.32$, $F(1, 24) = 5.80$, $p < .05$), a finding consistent with the adaptive network models but not HAM models. In condition 1, I₂ was perceived as a better predictor of moistness than I₁ ($M_{F_{3I_1}} = 6.25$, $M_{F_{3I_2}} = 7.68$, $F(1, 24) = 4.33$, $p < .05$), a finding consistent with both classes of models. In condition 2, I₁ was perceived as a better predictor of chocolate flavor than I₂ ($M_{F_{3I_1}} = 7.10$, $M_{F_{3I_2}} = 5.90$, $F(1, 26) = 8.20$, $p < .05$), a finding consistent with both classes of models. In condition 2, I₂ was perceived as a better predictor of moistness than I₁ ($M_{F_{3I_1}} = 5.77$, $M_{F_{3I_2}} = 6.67$, $F(1, 26) = 3.59$, $p < .07$), a

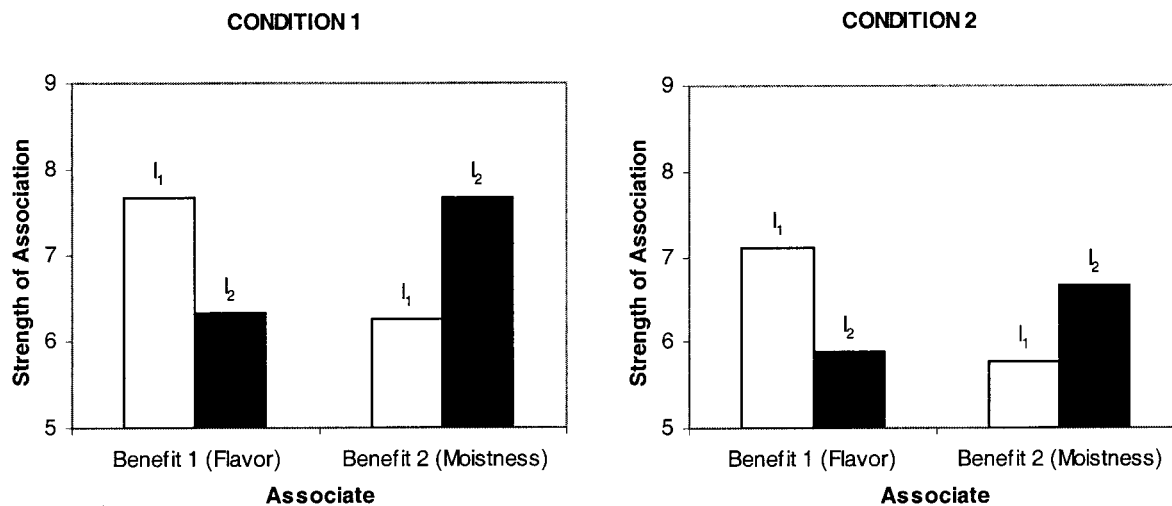
finding consistent with the adaptive network models but not the HAM models.

Discussion

In study 2, subjects were encouraged to focus on two consumption benefits during the tasting of the product samples. Unlike study 1, cue interaction was found for both outcomes in study 2. These results confirm the hypothesis that when consumers are encouraged to focus on more than one outcome during learning, they are able to engage in

FIGURE 3

STRENGTH OF ASSOCIATION BETWEEN INGREDIENT BRAND NAMES AND BENEFITS IN STUDY 2



adaptive learning about more than one outcome. The results disconfirm the hypothesis that the adaptive learning system is constrained to learning about a single outcome.

Studies 1 and 2 support the hypothesis that adaptive learning is found only when an outcome is the focus of prediction. Yet, these studies do not provide insight into the factors that make an outcome the focus of prediction. One possibility is that active, adaptive learning accompanies any explicit learning goal, whereas HAM learning accompanies any implicit learning goal. In other words, consumers engage in adaptive learning when they pay attention to an outcome but engage in HAM learning when they do not pay attention to an outcome. A second possibility is that adaptive learning accompanies any explicit learning goal about an outcome that is significant or worthy of prediction. In other words, an outcome must (1) receive attention and (2) be significant enough to become the focus of prediction during learning. These competing hypotheses are the subject of study 3.

STUDY 3

In study 3, we tested the hypotheses that adaptive learning only occurs when consumers are motivated to predict an outcome and that the motivation to predict an outcome is not a consequence of simply paying attention. We believe that consumers will be most motivated to learn to predict outcomes that they perceive to have hedonic relevance—that is, outcomes that are seen to have rewarding or punishing implications. Associates that are characteristics of the consumption experience itself, such as taste experiences, will often be seen to have direct hedonic relevance, and paying explicit attention to them will likely be sufficient for them to become the focus of prediction during learning. In con-

trast, associates that are characteristics of the product will not have any direct hedonic relevance, and merely paying attention to them will not be sufficient to make them become the focus of prediction during learning. For those associates to become significant enough to be the focus of prediction, they would have to be linked to hedonically relevant characteristics of the consumption experience that are the focus of prediction. Thus, we predict that paying explicit attention to an outcome is not sufficient for adaptive learning to occur but that outcomes also need to be hedonically relevant, either directly or indirectly, by being a proxy for another outcome that is the focus of prediction.

To test whether paying explicit attention was a sufficient condition for adaptive learning, we needed to manipulate the perceived hedonic relevance of an associate without adding confounding factors. One possible solution was to add a third associate that was not directly hedonically relevant but that could be varied with respect to indirect perceived hedonic relevance. Our third associate was the type of cocoa used in the cake mix. The type of cocoa in a cake mix (natural or Dutch cocoa) should have little hedonic relevance when the subjects first come to the experiment. Still, it is possible to make type of cocoa hedonically relevant indirectly, without changing other aspects of the stimuli. This could be done by making sure that the hedonically irrelevant associate (type of cocoa) was perfectly correlated with another, hedonically relevant associate (chocolate flavor) and by instructing subjects to focus either on that hedonically relevant associate or on a third associate (moistness). Only in the first case would subjects be sufficiently motivated to make the associate that was initially hedonically less relevant (type of cocoa) the focus of prediction during learning. Thus, when chocolate flavor is itself the focus of prediction,

type of cocoa would function as a proxy for a focal hedonically relevant outcome and would itself be learned adaptively. When chocolate flavor is not the focus of prediction during learning, type of cocoa would be, at most, a proxy for a nonfocal outcome, and HAM learning would occur.

Design and Predictions

In study 3, we added the type of cocoa ingredient information to the labels used to identify the cake samples, using a design otherwise identical to study 1 (see App. B). Mild chocolate flavor samples were labeled as having natural cocoa, and strong chocolate flavor samples were labeled as having Dutch cocoa. Like study 1, one-half of the subjects were asked to assess the moistness of the samples and the other half of the subjects were asked to assess the flavor of the samples. In addition, all subjects were asked to indicate whether the sample had Dutch cocoa or natural cocoa. Thus, one-half of the subjects were asked to rate the presence or absence of the associate that had little direct hedonic relevance (Dutch cocoa) and the presence or absence of the related, hedonically relevant associate (strong chocolate flavor). The other half of the subjects were asked to rate the presence or absence of the associate that had little direct hedonic relevance (Dutch cocoa) and the presence or absence of the unrelated, hedonically relevant associate (high moistness). The ingredient labels were counterbalanced.

At test, subjects were told that a muffin manufacturer believed that people preferred a product with Dutch cocoa, and, based on their experiences, they should indicate which ingredient brand they most associated with Dutch cocoa. Subjects were asked to indicate, "Which ingredient (e.g., Baker's Blend chocolate syrup vs. Silk'n Morsels chocolate bits) is made with real Dutch cocoa?" by circling the appropriate brand name. If paying explicit attention to an outcome and recording its level are sufficient to make an outcome the object of an adaptive learning process, then subjects should show cue interaction in both conditions. If, in addition to paying attention to an outcome, the outcome must be hedonically relevant, either directly or indirectly, in order for the outcome to become the object of an adaptive learning process, then subjects should show cue interaction only when they were asked to rate chocolate flavor of the samples. In the moistness rating condition, chocolate flavor should remain hedonically irrelevant and should not become the focus of predictive learning; hence, no cue interaction should be observed.

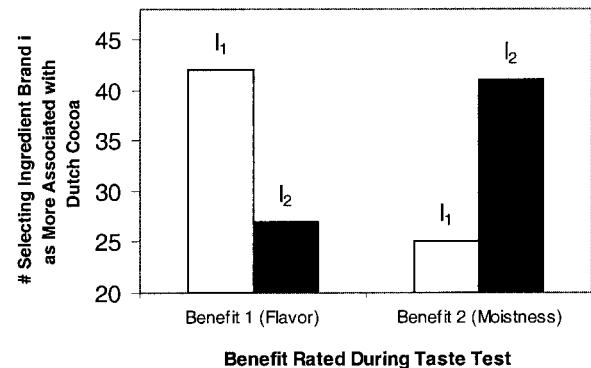
Results

One hundred thirty-five subjects from an introductory marketing class were awarded extra credit to participate in the study. The key dependent measure was the subjects' choice of I_1 or I_2 as the brand they associated most with Dutch cocoa. Figure 4 shows the number of subjects selecting I_1 or I_2 as the brand made with Dutch cocoa when they had also judged flavor or moistness. The overall test for an interaction between benefit judged during learning

FIGURE 4

RESULTS OF STUDY 3

STRENGTH OF ASSOCIATION BETWEEN INGREDIENT BRAND NAMES I_1 / I_2 AND ATTRIBUTE (DUTCH COCOA)



(flavor, moistness) and the brand chosen was significant ($\chi^2(1) = 7.13, p < .05$). Subjects who judged flavor intensity and the type of cocoa attribute during the learning phase were more likely to select I_1 as the brand made with Dutch cocoa ($N_{F_3I_1} = 42, N_{F_3I_2} = 27, \chi^2(1) = 3.26, p = .07$), a finding consistent with the adaptive network models but not the HAM models. Subjects who judged moistness and the type of cocoa attribute during the learning phase were more likely to select I_2 as the brand made with Dutch cocoa ($N_{F_3I_1} = 25, N_{F_3I_2} = 41, \chi^2(1) = 3.87, p < .05$), a finding consistent with the HAM models but not the adaptive network models.

Discussion

Study 3 supports the view that adaptive learning only occurs when consumers are motivated to make an outcome the focus of prediction during learning. When an associate that had little direct hedonic relevance (type of cocoa) became a clear proxy of a focal, hedonically relevant outcome (chocolate flavor), it became the focus of prediction and associations were learned adaptively, in accordance with the adaptive network models. When an associate that had little direct hedonic relevance (type of cocoa) was not a proxy of a focal, hedonically relevant outcome (chocolate flavor), the originally nonvalenced attribute did not become the focus of prediction and associations were learned in accordance with the HAM learning models. Thus, this study provides further evidence that adaptive learning and cue interaction require that associates become the focus of prediction during learning. Adaptive learning does not occur any time that associates are consciously attended and considered. Instead, adaptive learning occurs when an attended outcome is motivationally significant enough to make it the focus of prediction.

The presence of cue interaction when consumers focus on a hedonically relevant associate as an outcome to be predicted, a finding consistent with adaptive network models, and of cue independence when consumers do not try to learn to predict an associate, a finding consistent with HAM models, suggests that consumers use two different learning processes to form product associations. However, the presence or absence of cue interaction should not be the only difference between the two ways of learning associations that can be used to judge products. For example, in our description of the two processes, we proposed that judging products using adaptive associations is a forward-looking process and that judging products using HAM associations is a backward-looking process. If this is indeed the case, then it should be possible to change judgments by manipulating the evaluation strategies at the time of judgment, which is the main purpose of study 4.

STUDY 4

In study 4, we directly assess the forward-looking versus backward-looking nature of the two processes by instructing some subjects to try to recall individual experiences prior to predicting the quality of products. We also make a number of small changes in the procedure to handle potential criticisms of the first three studies. First, we use a nonexperiential outcome to ensure that adaptive learning is not limited to benefits that are directly experienced. Second, we directly measure the strength of associations during an initial learning phase to ensure that assumptions about learning of F_1 -outcome and F_2 -outcome associations are appropriate. In the first studies, some parameter configurations exist that could theoretically allow some versions of adaptive network models, such as Pearce's (1994) configural model, to account for the pattern of results (if the specification of the model is changed to allow learning parameters to vary for every cue-outcome combination). Even though these parameter values would have to be unusually low in some conditions, and quite varied across the first three studies, showing that HAM learning occurs even when the family brand name association to the target outcome reached an asymptote prior to the presentation of family/ingredient brand trials would provide irrefutable evidence against a complete explanation in terms of the adaptive network models. In addition to asymptotic family brand learning, irrefutable evidence in support of the two-process theory could also be obtained by showing two different types of learning after an identical set of learning trials. Logically, no single model can ever predict different learning results between two conditions if the learning part of the experiment is identical for both conditions. Thus, this study strongly challenges the ability of any one-process theory to explain how consumers learn product associations and use them to make predictions and evaluations.

Design and Procedure

The study was a two-cell, within-subject design (ingredient brand name that uniquely predicts the target benefit) with a between-subject instruction manipulation (no instruction, backward-looking instruction) and a brand name counterbalancing factor. The basic design was similar to study 1 but included an initial learning phase in which the F_1 and F_2 associations with the focal outcome were learned up to asymptote. The symbols F_1 and F_2 represented family brand names (e.g., Delight, Buon Chocolate). Outcome information was limited to a single benefit, mild or strong chocolate flavor.

Subjects participated in a computerized learning experiment in groups of up to 12 persons. On the first computer screen, subjects were instructed that they were to learn to predict the chocolate flavor of cakes. On the next four screens, they were presented with a brand name and asked to anticipate the flavor intensity of the brand by selecting a scale item labeled "mild chocolate flavor," "average chocolate flavor," or "strong chocolate flavor." Then, subjects were given feedback on the flavor of the sample using the schedule F_1 /mild chocolate flavor, F_2 /strong chocolate flavor, F_1 /mild chocolate flavor, and F_2 /strong chocolate flavor. After receiving the four product descriptions, subjects were asked to rate explicitly the likelihood that a product carrying the F_1 brand would have a strong chocolate flavor using an 11-point scale, with end points labeled "mild chocolate flavor" and "strong chocolate flavor." The same question was asked about the F_2 brand. If subjects did not select one of the two lowest scale values for the F_1 brand and the highest two scale values for the F_2 brand, they were given four additional trials. This procedure was repeated until subjects selected the appropriate extreme scale values or until they perfectly predicted the outcomes on the four most recent trials.

After the subjects reached asymptotic learning for both family brands, they were presented with a series of new products. The sequence they experienced was F_1 , F_1I_1 , F_2 , F_2I_2 , F_2I_2 . The symbols I_1 and I_2 represented ingredient brand names (e.g., Baker's Blend Syrup, Silk'n Morsels). All but the F_1 product were described as having a strong chocolate flavor. Next, subjects were asked to rate the likelihood that a new brand of muffin with each ingredient brand (F_3I_1 and F_3I_2) would have a strong chocolate flavor, using a 100-point scale with end points labeled "The muffin with Baker's Blend Syrup will have the strongest chocolate flavor" and "The muffin with Silk'n Morsels will have the strongest chocolate flavor." Right before answering this question, the subjects in the backward-looking condition were asked to think back and recall the product descriptions that they had seen in order to determine how likely the product would be to have a strong chocolate flavor. The other half of the subjects were not given the backward-looking instruction. Thus, the manipulation took place after learning but before the collection of the dependent measures. Therefore, no learning parameter differences could exist between the con-

ditions. We also note that there was a $F_2, F_2I_2, F_2I_2, F_1, F_1I_1$ order counterbalancing sequence for the stimuli.

We assume that, because adaptive learning is a more focused, motivationally guided process, consumers will spontaneously use adaptive associations to make predictions whenever adaptive associations are available. Because HAM learning always associates all elements of an experience, we assume that HAM associations will only guide predictions when either no adaptive associations are available or external factors (such as explicit instructions) encourage the use of HAM associations over adaptive associations. Thus, we predict that cue independence will occur when consumers are asked to retrospect and that cue interaction will occur when consumers are not asked to retrospect. According to the HAM models, I_2 should have a stronger association with strong chocolate flavor than does I_1 , because I_2 was paired with strong chocolate flavor two times, whereas I_1 was paired with the strong chocolate flavor once (see predicted strength of association graph in Fig. 1). According to adaptive network models, the I_1 cue will form a stronger association with strong chocolate flavor than the I_2 cue because the ingredient brand name uniquely identifies strong chocolate flavor in the first product line but is a redundant predictor for strong chocolate flavor in the second product line (see predicted strength of association graph in Fig. 1).

Results

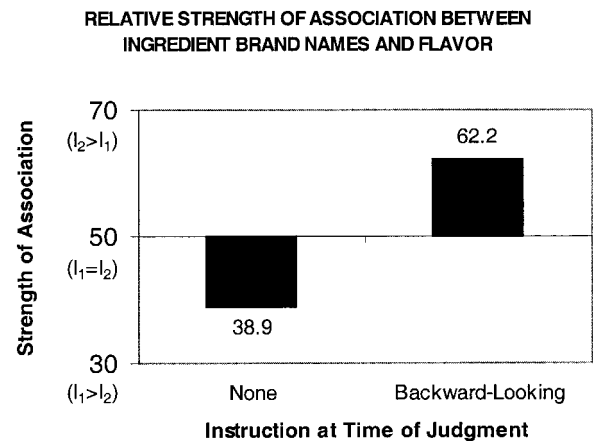
Forty subjects from an introductory marketing class were awarded extra credit to participate in the study. The expectation about the intensity of the chocolate flavor of the ingredient-branded muffin was used as an indicator of the association strength between each ingredient brand name and flavor. There was no influence of the counterbalancing factor ($F(1, 37) = 0.04, p > .10$).

Manipulation Check. All of the subjects reached asymptotic learning. Fifty-five percent of the subjects judged an F_1 product to be very unlikely to have a strong chocolate flavor and an F_2 product to be very likely to have a strong chocolate flavor. The remaining 45 percent of the subjects were not as extreme in their judgments but did correctly predict the flavor outcome in their last four asymptotic learning trials. There was no interaction between this indicator of asymptotic learning and the dependent measure ($F(1, 37) = 0.34, p > .10$), implying that differences in first-phase learning of the family brand associations could not account for the results on the dependent measure and again suggesting that all subjects had learned to asymptote.

Associations with Ingredient Brand Names. Figure 5 shows the degree of association between I_1 and I_2 , with chocolate flavor by condition. The backward-looking instruction significantly influenced the reported association between the ingredient brand names and moisture ($F(1, 37) = 5.31, p < .05$). In the no instruction condition, I_1 was perceived as a better predictor of chocolate flavor than I_2 ($M = 38.9$, one-tail test against the scale indifference

FIGURE 5

RESULTS OF STUDY 4



point of 50, $F(1, 37) = 2.89, p < .05$), a finding consistent with adaptive network models, but not the HAM models. In the backward-looking condition, I_2 was perceived as a better predictor of chocolate flavor than I_1 ($M = 62.2$, one-tail test against the scale indifference point of 50, $F(1, 37) = 3.20, p < .05$), a finding consistent with the HAM models, but not the adaptive network models.

Discussion

Results in study 4 provide direct support for a critical process-assumption of our two-process theory. In a situation in which all subjects should have formed adaptive associations in addition to HAM associations, subjects using a forward-looking judgment process showed cue interaction, but subjects using a backward-looking judgment process showed cue independence. In addition, the outcome was not experiential, suggesting adaptive learning can occur in contexts that have nonexperiential outcomes as the focus of prediction. Thus, our data again suggest that both classes of models have merit.

This finding conceptually replicates van Osselaer and Alba's (2000) experiment 2A. However, their experiment was designed merely to disconfirm a purely attentional explanation of the blocking effect. Unsurprisingly, it had design limitations that made it unsuitable as a test of the backward-looking, cue-independent learning hypothesis. First, van Osselaer and Alba asked subjects to write down the frequencies of co-occurrence between the to-be-blocked cue and the outcome. This recording task highlighted the fact that the frequencies for the to-be-blocked cue were as high as those for the blocking cue. Van Osselaer and Alba (2000) mention that this procedure might easily lead to a consistency-in-responding bias in which subjects try to avoid appearing inconsistent to the experimenter by assigning no weight to the blocked cue in the immediately following

prediction task. In our experiment, we do not ask subjects to compute and record frequencies, and subjects give only one response, the predictive response. Thus, no consistency in responding bias can occur. Second, in van Osselaer and Alba's (2000) experiment, the backward-looking instruction merely reduced the blocking effect, which could result from several factors such as the extra delay between learning and predictive judgment in the backward-looking condition. In our experiment, the difference between adaptive learning and HAM learning in forward and backward learning contexts is supported with a true judgment reversal that cannot be explained by a mere weakening of adaptive associations over time.

The results also disconfirm the possibility that all of the results in the first three studies can be explained by an adaptive network model that assumes very low learning parameters for the family brands in some conditions. Algebraic analysis and systematic parameter search simulations of study 4 show that there exists no set of parameter settings that can make the LMS or Pearce models fit the cue-independent result under asymptotic learning of family brand associations. In addition, both groups in study 4 received exactly the same learning treatment. Therefore, both groups learned the same way, with exactly the same set of learning parameters. Logically, no model can predict two different results with a single parameter setting. Thus, both the asymptotic learning aspect and the fact that learning itself was not manipulated bolster the claim that the results cannot be the result of any one process.

GENERAL DISCUSSION

Understanding how consumers learn product associations is important for researchers seeking to understand how consumers evaluate products and make choices. Research in psychology and consumer behavior has recently suggested that the associations people use to make product evaluations, predictions, and choices are the result of an adaptive learning process characterized by cue interaction (e.g., Janiszewski and van Osselaer 2000; Pearce 1994; Shanks et al. 1996). The findings of cue interaction in the studies suggested that the HAM models that have traditionally guided consumer research on brand associations are either incomplete or incorrect. We find that consumers have two associative learning processes at their disposal that can guide product evaluations, predictions, and choices.

We have also proposed a more detailed theory specifying the qualitative differences between the two processes and have started to test it empirically. Our data support findings by van Osselaer and Alba (2000) showing that cue interaction is the result of a forward-looking process and extend them significantly by showing that the other, cue-independent process is inherently backward looking. In addition to showing that there are two learning systems that can be used to make predictive judgments and illustrating that the second process is cue independent and backward looking, we investigate factors that determine which process guides behavior. Our results show that adaptive learning drives out-

come predictions only when the outcome is the focus of prediction during learning. We find that merely paying attention to an outcome is not sufficient for that outcome to become the focus of prediction during learning. Study 3 shows that even when an outcome is explicitly attended to, HAM learning will drive predictions unless the outcome is perceived to be hedonically relevant. Thus, it seems that adaptive learning is restricted to outcomes that are motivationally significant.

It is important to note that the predictions tested in our studies are not specific to the four models discussed here. These models are merely representatives of their respective classes of models, chosen because they are the most current, prototypical, and empirically supported in their respective literatures. For example, adaptive network models that we have not discussed, including Attention Learning COVERing Map (ALCOVE; Kruschke 1992), Attention to Distinctive Input (ADIT; Kruschke 1996), and Rapid Attention Shifts 'N' Learning (RASHNL; Kruschke and Johansen 1999), use learning rules that have the same basic characteristics of error minimization and additivity that characterize the configural and LMS models. Thus, the strong theory tests in our studies investigate the predictions of two classes of models, as opposed to the predictions of specific models.

It is also important to recognize that the proposed two-process theory has the potential to provide insights into why different learning goals and/or outcomes promote HAM learning or adaptive learning in a variety of contexts. For example, learning can be incidental or intentional, nonexperiential or experiential, about nonvalenced attributes or valenced benefits, and about beliefs or attitudes. To the extent that learning involves outcomes that are more motivationally significant because they are very salient or more closely related to the actual consumption experience, adaptive learning should occur.

Finally, it is important to note that the findings presented here have important managerial implications. Whenever consumers use the HAM learning system, brand equity will to a large extent depend on the frequency with which brand names are accompanied by positive consumption experiences. This implies, for example, that adding a well-known ingredient brand name to a well-known product without changing its quality will strengthen ingredient brand associations to quality. In contrast, whenever consumers use the adaptive learning system, the equity of a brand name will depend on what consumers learn about other features of the product. For example, if the adaptive system governs consumers' behavior, adding a well-known ingredient brand name to a well-known product without changing its quality will often weaken the ingredient brand association to quality because of an overexpectation effect (Janiszewski and van Osselaer 2000).

Limitations and Future Research

Together with previous work by consumer researchers, these results provide insight into how consumer learning affects product evaluation and choice. However, the litera-

ture on consumer learning and evaluation is still small, and many questions remain unanswered. The empirical studies reported here cover only a few of the many testable implications of a learning theory that incorporates both HAM learning and adaptive learning. For example, we have not explored the validity of the supposed shortsightedness of the adaptive learning system, a system that changes association strengths just to increase the accuracy of the next prediction yet is insensitive to the particular learning path that led to the associations.

In addition, our theory is incomplete because it only makes predictions about some aspects of the two learning processes. For example, it does not specify whether the two learning systems are configural or elemental, which is important because configural processes should generally lead to weaker transfers of brand associations within and between brand portfolios. That is, if learning is configural, it should be more product specific, and learning about one product should have less influence on evaluations of other products that share only some of the base product's features. In the adaptive learning literature, Shanks and his colleagues have recently found strong configural learning effects (Lopez et al. 1998; Shanks, Charles, et al. 1998; Shanks, Darby, and Charles 1998). That is, learning cannot be described solely in terms of the elements (or cues) that make up a stimulus, as in the LMS, ADIT, and RASHNL models (Gluck and Bower 1988; Kruschke 1996; Kruschke and Johansen 1999). Instead, learning models need to include representations of the stimulus as a whole, as in configural models such as Pearce's (1994) configural model and Kruschke's (1992) ALCOVE model, in order to be able to explain both elemental cue interaction effects and configural effects. However, some catastrophic interference experiments showing configural effects suggest that such hybrid models that try to explain both elemental and configural effects are not configural enough (Shanks, Charles, et al. 1998; Shanks, Darby, and Charles 1998). In contrast, purely configural models, in which only the presented configuration is activated, can never account for cue interaction. Thus, it seems that elemental and configural learning both exist (cf. Williams et al. 1994).

We speculate that elemental and configural effects are not the result of one single process and that any attempt to create one model describing both is doomed to fail. Instead, we believe that elemental learning effects are the result of adaptive learning and that configural learning effects are the result of HAM learning. This speculation is motivated by our hypothesis that adaptive learning is a more focused and selective process that identifies specific stimulus elements as outcomes to be predicted and other specific elements as cues used to predict. In contrast, we describe HAM learning as a more general process that stores information from the whole stimulus. Configural adaptive learning models such as Pearce's (1994) add configurality to a basic LMS model in an attempt to describe both processes. It is possible that adaptive learning is well described by a very simple elemental model. Pearce's configural addition may be useful

as the foundation of a cue-independent, configural HAM learning process. In particular, we believe that a model that uses Pearce's general network structure and configural activation process but replaces the LMS-type configural node-outcome learning rule (App. A, Eq. A11) by a simple Hebbian updating rule (App. A, Eq. A5) may be promising as a model of HAM learning.

Another area for future exploration is whether the adaptive learning process is sequential or parallel. Most models outlined in this article are essentially parallel models, because each learning experience leads to the simultaneous updating of more than one association. However, many characteristics of human judgment can be explained in terms of selective sequential hypothesis testing (Sanbonmatsu et al. 1998). According to Sanbonmatsu and his colleagues, human judgment is often not based on the simultaneous investigation of several different hypotheses but on a selective top-down process in which one hypothesis is tested at a time and hypothesis testing is halted as soon as one hypothesis provides a sufficient (but not necessarily optimal) answer to the question at hand. One of the phenomena easily explained by such a process is the blocking phenomenon. If consumers first confirm the hypothesis that the first cue (e.g., a family brand name) predicts the outcome, hypothesis testing stops and the predictive value of an additional cue (e.g., an ingredient brand name) is never assessed. Thus, learning about the predictive value of the additional cue is blocked. Although it is not clear that all cue interaction phenomena outlined here and elsewhere can be explained in terms of a selective, sequential hypothesis testing process (see Janiszewski and van Osselaer [2000] for a wide range of cue interaction phenomena), it is possible that updating in the adaptive learning system is sequential instead of parallel (van Osselaer and Alba 2000).

Recently, Kruschke and Johansen (1999) proposed a parallel model with an attention allocation function that speeds updating for one or a few cues while slowing updating for the remaining cues. Thus, this model seems to provide a compromise between a parallel process that has similar updating rates for all cues and a purely sequential system that has zero updating rates for all but one cue. An empirical test of these three processes (equal parallel updating, selective parallel updating, and selective sequential updating) would provide important insight into the exact nature of adaptive processing. On a more practical level, such a test would provide insight into the specificity of consumer learning, with implications for the optimal stimulus complexity in marketing communications used in adaptive learning contexts. In addition to providing an impetus for specifying and testing the selective hypothesis testing process, the experiments and theory presented here might help to clarify (1) the exact nature of the other type of processing in Sanbonmatsu et al.'s (1998) framework comparative processing and (2) when each type of processing governs consumer predictions and evaluations. We believe that there would be significant scientific benefit in the development of a formal model of sequential hypothesis testing (perhaps using the

RULE-plus-EXception [RULEX] model by Nosofsky, Palmeri, and McKinley [1994] as a starting point). In our experience, working with formal models has helped us to specify the exact nature of processes; to find, through thought experiments and simulations, conditions that allow clear falsification of a model; and to make new predictions about unknown phenomena (see Smith [1996] and Smith and DeCoster [1998] for a similar argument).

In addition to the dual-process theory by Sanbonmatsu et al. (1998), it would also be of interest to investigate the similarities and differences between our theory and another connectionist dual-process theory by Smith and DeCoster (1999). They discern two types of processing, associative and rule-based processing. Associative processing makes use of a slow learning system and rule-based processing makes use of both a fast and a slow learning system. Interestingly, these processes and learning systems do not seem to map directly onto our two processes and learning systems. In fact, the learning model used for the slow learning system, described by Smith and DeCoster (1998) and developed by McClelland and Rumelhart (1986), uses the same Delta learning rule as the adaptive learning models presented here. Smith and DeCoster (1999) are not very specific about the exact nature of their rule-based process. However, it is defined as a symbolic reasoning process rather than an associative process. Thus, it is possible that there are at least three learning processes involved in consumer prediction and evaluation—a slow and low-effort HAM-type process, a faster and higher-effort but still associative adaptive process, and an even faster and more resource-intensive symbolic reasoning process that can come up with abstract rules. This would be consistent with other recent findings in the learning literature showing that, in addition to an adaptive associative process, there might also be a rule-based process (e.g., Erickson and Kruschke 1998; Shanks and Darby 1998). In sum, there are interesting similarities and differences between our theory and these other theories. What most distinguishes ours is the inclusion of a cue-independent, backward-looking process used to make predictions and evaluations, plus our findings of moderators determining when which process will govern responses.

This leads us to the issue of the interaction between the two learning systems. Gilbert (1999) describes four ways in which two processes might interact. One possibility is selective combination, in which, on any occasion, only one system is active. A second possibility is competitive combination, in which both systems always produce an output, but only the strongest has any influence on behavior. Third, it is possible that combination is consolidative, in which both systems are active and both systems' output shows up in behavior. Fourth, combination might be corrective, in which one system is active by default and the other system is only activated when the first system fails and sensitively corrects the first. We believe the adaptive and HAM systems combine in a fifth way. Study 4 seems to support our assumptions about how the adaptive and HAM learning systems interact. Because HAM learning does not require con-

sumers actively to focus their learning on predicting a specific outcome, we expect that HAM learning is always active. That is, consumers always store and cross-reference all stimulus elements. In contrast, adaptive learning may be restricted to a narrower set of circumstances, because adaptive associations are only formed between product cues and outcomes that consumers are motivated to learn to predict. Thus, we believe that HAM associations are always formed but are overpowered by adaptive associations whenever the latter are available. This overpowering could be all or nothing or could take the form of a weighted sum, with either the intensity or weight of the adaptive system's output being much stronger than the HAM system's output. We speculate that it is not the case that only one process can be active at the same time (in contrast to selective combination), that it is not true that both processes are always active (in contrast to competitive or consolidative combination), and that one process is not sensitively correcting the output of the other process (in contrast to corrective combination).

In sum, the studies presented here reconcile two streams of research by showing that consumers have two learning systems at their disposal that can be used to make predictions about products. The studies investigate two critical aspects of these two systems, cue interaction versus cue independence and being forward looking versus backward looking, and several moderators that determine which system guides responses. However, many other differences may exist between the learning systems, and many other moderators may determine when each system drives consumer behavior.

APPENDIX A MODELS

In Appendix A, we describe the four learning models introduced in the first part of the article.

THE DA MODEL

The DA model uses a simple Hebbian or "coincident activation" learning rule (Janiszewski and van Osselaer 2000). In the DA model, the change in the association strength (Δs_{ij}) between two nodes i and j in one learning trial is given by

$$\Delta s_{ij} = \beta * a_i * q_j, \quad (\text{A1})$$

where β ($0 < \beta < 1$) is a learning rate parameter, a_i ($a_i = 0, 1$) is the activation level of cue i (i.e., whether the cue is present or not), and q_j is the experienced level of associate j (e.g., the value of the associate). Thus, the association between a cue and an associate is strengthened every time that the cue is presented with the associate and this learning is independent of the associations between other cues and the same associate. Finally, the level of activation of the associate j , expressed as o_j , is an additive function of incoming activation from all cues, or

$$o_j = \sum_{i=1, n} (a_i * s_{ij}). \quad (\text{A2})$$

THE ACT-R MODEL

According to the ACT-R model (Anderson 1993; Anderson and Lebiere 1998), declarative knowledge is represented as a network of concept nodes connected by associations that are strengthened each time two events co-occur. The ACT-R model predicts that the activation of an outcome node o_j is a function of that node's base level of activation (b_j) and of incoming activation of the outcome node j from available cueing nodes i (s_{ij}), weighted for their activation level (a_i), or

$$o_j = b_j + \sum_{i=1, n} (a_i * s_{ij}). \quad (\text{A3})$$

The ACT-R model proposes association strengths (s_{ij}) are a logarithmic function of (1) the association strength at the beginning of the learning session (r_{ij}^*); (2) a positive weighting constant (w); (3) the empirical ratio, e_{ij} , of the conditional probability of outcome node j 's presence given cue node i 's presence divided by the base rate of outcome node j 's presence; and (4) the frequency [$F(i)$] of the presentation of cue i . More formally, the associative strength between cue i and outcome j is

$$s_{ij} = \log(r_{ij}), \quad (\text{A4})$$

where

$$r_{ij} = [r_{ij}^* * w + e_{ij} * F(i)] / [w + F(i)], \quad (\text{A5})$$

and where

$$e_{ij} = P(j|i) / P(j). \quad (\text{A6})$$

In this model, it should be noted that the numerator and denominator include the common term [$w + F(i)$] but that the numerator includes a r_{ij}^* adjustment to the weighting constant (w) and an e_{ij} adjustment to the frequency of presentation [$F(i)$] term. First, it is important to recognize that without these adjustments, the ratio would take on the value of one, the log of the ratio would take on the value zero, and association strength between cue node i and outcome node j would be zero. Second, including the r_{ij}^* adjustment to the weighting constant (w) allows the model to represent prior learning (r_{ij}^* takes on a value of one when there is no prior learning and a value different from one when there is prior learning). Third, including the e_{ij} adjustment allows the model to increase association strengths for cues that are more strongly correlated with the outcome.

THE LMS MODEL

Identical to the DA model, the LMS adaptive network model (LMS model) assumes the activation of associate j

is equal to the sum of the incoming activation from all activated cues i (Eq. A2). In addition, the LMS model posits an error-reduction property, which says that the updating of association strengths is a function of the error between a predicted outcome and feedback about the outcome that is actually experienced (e.g., Gluck and Bower 1988; Janiszewski and van Osselaer 2000; McClelland and Rumelhart 1986). This results in the following learning rule:

$$\Delta s_{ij} = \beta * a_i * (q_j - o_j), \quad (\text{A7})$$

where Δs_{ij} is the change in the association from i to j , β ($0 < \beta < 1$) is a learning rate parameter, a_i ($a_i = 0, 1$) is the activation on input node i , q_j is the experienced outcome (e.g., whether a benefit is present), and o_j (defined in Eq. A2) is the expectation about whether the outcome will be present. Together, the additivity and error-reduction properties lead to a third property, cue interaction. That is, updating of the association between one cue and an associate depends on the presence of other cues in the same learning trial.

PEARCE'S (1994) CONFIGURAL MODEL

Pearce's (1994) configural model extends the LMS model to include configural effects and an attentional limitation. The configural model consists of three layers of connections between four layers of nodes. First, the presence or absence of a number of cues i is represented by a layer of input nodes i that are activated if cue i is present and not activated if cue i is absent. The input layer is connected to an equal number of hidden nodes i representing each input node's internal activation. This internal activation is a function of the activation of the corresponding input node and of the presence of other cues, such that activation of one cue's internal activation node is reduced when other cues are present. Thus, the first layer of connections performs an attentional function, reducing attention to any one cue with each additional cue that is present. Formally, the activation, a_i , of hidden node i if cue i is present is given by the following equation:

$$a_i = 1/\sqrt{n}, \quad (\text{A8})$$

where n is the total number of cues present in the stimulus pattern.

Cue-specific internal activation is then fed through to a second layer of hidden nodes representing whole stimulus configurations (i.e., if consumers encounter stimuli F_1 and F_1I_1 , two different configural nodes are created). Formally, the activation, c_j , of configural node j is given by

$$c_j = [\sum_{i=1, n} (a_i * s_{ij})]^2, \quad (\text{A9})$$

where a_i is the activation level of the internal activation node of cue i and where s_{ij} is the strength of the association between the internal activation node of cue i and configural

node j . Association strengths between cues' internal activation hidden nodes and configural hidden nodes are the result of a one-shot learning process. The first time a stimulus is encountered, the association strengths between the internal activation nodes of its elements and the corresponding configural node are set to be equal to the activation levels of the internal activation nodes. Thus, the mapping from internal activation nodes to configural nodes remains constant.

Finally, the activation of the configural nodes is fed forward from the configural nodes to one or more outcome nodes. As in the other models, activation, o_k , of each outcome node k is an additive function of incoming activations. Each incoming activation is determined by the activation level (c_j) of the configural node and by the strength of the association between the configural node and the outcome node (s_{jk}):

$$o_k = \sum_{j=1, m} (c_j * s_{jk}). \tag{A10}$$

In this model, the configural node-outcome associations s_{jk} are updated (learned) according to the following rule:

$$\Delta s_{jk} = \beta * c_j * (q_k - o_k), \tag{A11}$$

where Δs_{jk} is the change in the association from configural node j to outcome node k , β ($0 < \beta < 1$) is a learning rate parameter, c_j is the activation of configural node j , q_k is the experienced outcome, and o_k (defined in Eq. A10) is the expected level of the outcome. On each trial, only the configural node-outcome associations of the most strongly active configural node are updated. This should always be the configural node representing the currently present stimulus pattern.

APPENDIX B DESIGN SUMMARIES

TABLE B1

STUDY 1 AND SECOND PHASE OF STUDY 4 (BENEFIT 1 ONLY)

Design			Operationalization of design		
Brand	Benefit 1	Benefit 2	Sample brand	Benefit 1	Benefit 2
L	-	-	Treats	Mild cocoa	Dry
F ₁	-	+	Delight	Mild cocoa	Moist
F ₁ I ₁	+	+	Delight w/ Baker's Blend Syrup	Strong cocoa	Moist
F ₂	+	-	Buon Chocolate	Strong cocoa	Dry
F ₂ I ₂	+	+	Buon Chocolate w/ Silk'n Morsels	Strong cocoa	Moist
F ₂ I ₂	+	+	Buon Chocolate w/ Silk'n Morsels	Strong cocoa	Moist

TABLE B2

STUDY 2

Condition 1					
Design			Operationalization of design		
Brand	Benefit 1	Benefit 2	Sample brand	Benefit 1	Benefit 2
L	-	-	Treats	Mild cocoa	Dry
F ₁	-	+	Delight	Mild cocoa	Moist
F ₁ I ₁	+	+	Delight w/ Baker's Blend Syrup	Strong cocoa	Moist
F ₂	+	-	Buon Chocolate	Strong cocoa	Dry
F ₂ I ₂	+	+	Buon Chocolate w/ Silk'n Morsels	Strong cocoa	Moist
F ₂ I ₂	+	+	Buon Chocolate w/ Silk'n Morsels	Strong cocoa	Moist
Condition 2					
Design			Operationalization of design		
Brand	Benefit 1	Benefit 2	Sample brand	Benefit 1	Benefit 2
L	-	-	Treats	Mild cocoa	Dry
F ₁	-	+	Delight	Mild cocoa	Moist
F ₁ I ₁	+	+	Delight w/ Baker's Blend Syrup	Strong cocoa	Moist
F ₁ I ₁	+	+	Delight w/ Baker's Blend Syrup	Strong cocoa	Moist
F ₂	+	-	Buon Chocolate	Strong cocoa	Dry
F ₂ I ₂	+	+	Buon Chocolate w/ Silk'n Morsels	Strong cocoa	Moist

TABLE B3
STUDY 3

Design				Operationalization of design			
Brand	Attribute	Benefit 1	Benefit 2	Sample brand	Attribute 1	Benefit 1	Benefit 2
L	–	–	–	Treats	Natural	Mild	Dry
F ₂	+	+	–	Buon Chocolate	Dutch	Strong	Dry
F ₂ I ₂	+	+	+	Buon Chocolate w/ Silk'n Morsels	Dutch	Strong	Moist
F ₂ I ₂	+	+	+	Buon Chocolate w/ Silk'n Morsels	Dutch	Strong	Moist
F ₁	–	–	+	Delight	Natural	Mild	Moist
F ₁ I ₁	+	+	+	Delight w/ Baker's Blend Syrup	Dutch	Strong	Moist

APPENDIX C

ILLUSTRATED MODEL PREDICTIONS

In Appendix C, we illustrate how the models can be used to make predictions about the association strength between I₁ and I₂ and the benefit chocolate flavor. We note that the same base design is used in all four studies (see App. B). We also note that the assumption of asymptotic learning of F₁ benefit and F₂ benefit associations in F-only trials, used in the predictions of the LMS and Pearce models, is confirmed in study 4.

THE DA MODEL

The DA model predicts that the difference between the I₁-outcome and I₂-outcome associations should be equal to $-\beta$. For example, in study 1, the strength of the association between I₁ (e.g., Baker's Blend) and benefit 1 (B₁, e.g., chocolate flavor) should be equal to $\beta * a_i * q_j = \beta * 1 * 1 = \beta$. The strength of the association between I₂ (e.g., Silk 'n Morsels) and B₁ should be equal to $2 * \beta * a_i * q_j = 2 * \beta * 1 * 1 = 2\beta$, because I₂ and B₁ were presented together twice. Thus, the difference in outcome activations in the F₃I₁ and F₃I₂ test trials is equal to $\beta - 2\beta = -\beta$.

THE ACT-R MODEL

The ACT-R model also predicts that the difference between the I₁-outcome and I₂-outcome associations should be negative. In our experiments, we keep e_{ij} constant, typically at 1.5 (i.e., $P(j|i)/P(j) = 1.0/.67$) for both ingredient brand associations. We keep the r_{ij}^* at 1 (i.e., both ingredient brands are novel). The strength of the association between I₁ and the flavor benefit should, then, be equal to

$$\begin{aligned}
 s_{I_1-B_1} &= \log \{ [r_{I_1-B_1}^* * w + e_{I_1-B_1} * F(I_1)] / [w + F(I_1)] \} \\
 &= \log [(1 * w + 1.5 * 1) / (w + 1)] \\
 &= \log [(w + 1.5) / (w + 1)].
 \end{aligned}$$

The strength of the association between I₂ and the flavor benefit should be equal to

$$\begin{aligned}
 s_{I_2-B_1} &= \log \{ [r_{I_2-B_1}^* * w + e_{I_2-B_1} * F(I_2)] / [w + F(I_2)] \} \\
 &= \log [(1 * w + 1.5 * 2) / (w + 2)] \\
 &= \log [(w + 3) / (w + 2)].
 \end{aligned}$$

Thus, the difference in outcome activation between the F₃I₁ and F₃I₂ test trials should be equal to $\log[(w + 1.5)/(w + 1)] - \log[(w + 3)/(w + 2)]$. This difference is negative for all positive values of w .

THE LMS MODEL

The LMS model predicts that the difference between the I₁-outcome and I₂-outcome associations should be positive. For example, suppose experiencing the mild chocolate flavor (B₁) of the Delight product (F₁) is represented as a negative outcome (q_j) of -1 . For predicted and experienced outcome to be equal, the outcome prediction on F₁-only trials has to be -1 . This implies that the asymptotic F₁-B₁ association strength also has to be -1 (using Eq. A2, $o_{B_1} = 1 * s_{F_1-B_1} + 0 * s_{F_2-B_1} + 0 * s_{I_1-B_1} + 0 * s_{I_2-B_1} = -1$). Suppose experiencing the strong chocolate flavor (B₁) on the next trial of Delight with Baker's Blend Syrup (F₁I₁) is represented as $+1$. To decrease the large prediction error (which should be equal to $q_{B_1} - o_{B_1} = 1 - [-1] = 2$), the I₁-B₁ association should be updated to

$$\begin{aligned}
 \Delta s_{I_1-B_1} &= \beta * a_{I_1} * (q_{B_1} - o_{B_1}) \\
 &= \beta * 1 * [1 - (-1)] = 2\beta,
 \end{aligned}$$

where β is a learning rate that cannot be negative (see Eq. A7). Thus, after the F₁I₁ trial, $s_{I_1-B_1}$ equals 2β . For the F₂ family brand name (e.g., Buon Chocolate), the positive experience can be represented as a $+1$, and $s_{F_2-B_1}$ should be 1. Experiencing the strong chocolate flavor (B₁) on the next trial of Buon Chocolate with Silk'n Morsels can also be

represented as a +1. Because $s_{F_2-B_1}$ is equal to 1, the prediction error on the F_2I_2 trials is zero, and there should be no updating of the I_2-B_1 association or

$$\begin{aligned}\Delta s_{I_2-B_1} &= \beta * a_{I_2} * (q_{B_1} - o_{B_1}) \\ &= \beta * 1 * (1 - 1) = 0.\end{aligned}$$

Thus, after the two F_2I_2 trials, $s_{I_2-B_1}$ equals zero. In sum, the difference in outcome activations in the F_3I_1 and F_3I_2 test trials should be positive and equal to 2β .

PEARCE'S (1994) CONFIGURAL MODEL

The Pearce (1994) model also predicts that the difference between the I_1 -outcome and I_2 -outcome associations should be positive. When F_1 -only is presented, the F_1 configural node should have an activation of

$$\begin{aligned}c_{F_1} &= [\sum_{i=1, n} (a_i * s_{ij})]^2 \\ &= (a_{F_1} * s_{F_1-F_1})^2 \\ &= (1 * 1)^2 = 1.\end{aligned}$$

The same is true for the F_2 configural mode. To minimize error when F_1 -only is presented, the asymptotic association between the F_1 configural node and benefit 1 (B_1) should be equal to -1 , so that $o_{B_1} = \sum_{j=1, m} (c_j * s_{jk}) = c_{F_1} * s_{F_1-B_1} = 1 * -1 = -1$. Next, on the F_1I_1 trial, internal activation of the F_1 and I_1 cue-specific nodes is equal to $1/\sqrt{2} = .7071$. The associations between the F_1 internal activation node and the F_1I_1 configural node, and between the I_1 internal activation node and the F_1I_1 configural node, should both also be equal to $.7071$. This leads to an activation of the F_1I_1 configural node of

$$\begin{aligned}c_{F_1I_1} &= [(a_{F_1} * s_{F_1-F_1I_1}) + (a_{I_1} * s_{I_1-F_1I_1})]^2 \\ &= [(.7071 * .7071) + (.7071 * .7071)]^2 = 1.\end{aligned}$$

It also leads to an activation of the F_1 configural node of

$$\begin{aligned}c_{F_1} &= [(a_{F_1} * s_{F_1-F_1}) + (a_{I_1} * s_{I_1-F_1})]^2 \\ &= [(.7071 * 1) + (.7071 * 0)]^2 = .5.\end{aligned}$$

Given that the F_1I_1 configural node has not yet established an association with benefit 1 and that the F_1 configural node has an association strength of -1 , the outcome activation on the F_1I_1 trial should be equal to

$$\begin{aligned}o_{B_1} &= (c_{F_1I_1} * s_{F_1I_1-B_1}) + (c_{F_1} * s_{F_1-B_1}) \\ &= (1 * 0) + (.5 * -1) = -.5.\end{aligned}$$

The $F_1I_1-B_1$ association should then be updated to be equal to

$$\begin{aligned}\Delta s_{F_1I_1-B_1} &= \beta * c_{F_1I_1} * (q_{B_1} - o_{B_1}) \\ &= \beta * 1 * [1 - (-.5)] = \beta * 1.5.\end{aligned}$$

At test, the F_3I_1 stimulus should activate the F_3I_1 configural node, which has no association to benefit 1, but should also partially activate the F_1I_1 configural node, at

$$\begin{aligned}c_{F_1I_1} &= [(a_{F_3} * s_{F_3-F_1I_1}) + (a_{I_1} * s_{I_1-F_1I_1})]^2 \\ &= [(.7071 * 0) + (.7071 * .7071)]^2 = .25.\end{aligned}$$

This should lead to an outcome activation on the F_3I_1 test trial of

$$\begin{aligned}o_{B_1} &= (c_{F_3I_1} * s_{F_3I_1-B_1}) + (c_{F_1I_1} * s_{F_1I_1-B_1}) \\ &= (1 * 0) + (.25 * \beta * 1.5) = .375\beta.\end{aligned}$$

To minimize error when F_2 -only is presented, the asymptotic association between the F_2 configural node and benefit 1 (B_1) should be equal to 1. Next, with regard to the activation of configural nodes, the situation is exactly parallel to that for the F_1I_1 stimulus. On the F_2I_2 trials, internal activation of the F_2 and I_2 cue-specific nodes is equal to $.7071$. The associations between the F_2 internal activation node and the F_2I_2 configural node, and between the I_2 internal activation node and the F_2I_2 configural node, are also equal to $.7071$, leading to an activation of the F_2I_2 configural node of 1 and of the F_2 configural node of $.5$. Because the F_2I_2 configural node has a zero association with benefit 1 and the F_2 configural node has an association strength of 1, the outcome activation on the first F_2I_2 trial is equal to

$$\begin{aligned}o_{B_1} &= (c_{F_2I_2} * s_{F_2I_2-B_1}) + (c_{F_2} * s_{F_2-B_1}) \\ &= (1 * 0) + (.5 * 1) = .5.\end{aligned}$$

The $F_2I_2-B_1$ association should be updated to

$$\begin{aligned}\Delta s_{F_2I_2-B_1} &= \beta * c_{F_2I_2} * (q_{B_1} - o_{B_1}) \\ &= \beta * 1 * (1 - .5) = \beta * .5.\end{aligned}$$

On the second F_2I_2 trial, the output activation of benefit 1 should then be equal to

$$\begin{aligned}o_{B_1} &= (c_{F_2I_2} * s_{F_2I_2-B_1}) + (c_{F_2} * s_{F_2-B_1}) \\ &= (1 * \beta * .5) + (.5 * 1) = (\beta * .5) + .5.\end{aligned}$$

This leads to further updating of the $F_2I_2-B_1$ association by

$$\begin{aligned}\Delta s_{F_2I_2-B_1} &= \beta * c_{F_2I_2} * (q_{B_1} - o_{B_1}) \\ &= \beta * 1 * [1 - (\beta * .5 + .5)] = .5\beta - .5\beta^2.\end{aligned}$$

Thus, the final strength of the F_2I_2 - B_1 association is equal to $(.5\beta) + (.5\beta - .5\beta^2) = \beta - .5\beta^2$. At test, the F_3I_2 stimulus should activate the F_3I_2 configural node completely and should activate the F_2I_2 configural node at .25. This leads to an outcome activation on the F_3I_2 test trial of

$$\begin{aligned}o_{B_1} &= (c_{F_3I_2} * s_{F_3I_2-B_1}) + (c_{F_2I_2} * s_{F_2I_2-B_1}) \\ &= (1 * 0) + [.25 * (\beta - .5\beta^2)] = .25\beta - .125\beta^2.\end{aligned}$$

Thus, the difference in outcome activations in the F_3I_1 and F_3I_2 test trials according to the Pearce model should be positive and equal to $.375\beta - (.25\beta - .125\beta^2) = .125\beta + .125\beta^2$.

Despite the fact that the four models make clear predictions in terms of outcome activations, we cannot make predictions about subjects' responses beyond the facts that the HAM models predict a negative difference between benefit 1 ratings of F_3I_1 and F_3I_2 and that the adaptive network models predict a positive difference. This is the case because the learning rates should be allowed to differ between models and because the function relating outcome activations to responses is not constrained beyond being monotonically positive. Thus, any learning rate could lead to any size, but not any direction, of response difference.

[Received March 2000. Revised February 2001. David Glen Mick served as editor, and Frank R. Kardes served as associate editor for this article.]

REFERENCES

- Alba, Joseph W., J. Wesley Hutchinson, and John G. Lynch, Jr. (1991), "Memory and Decision Making," in *Handbook of Consumer Behavior*, ed. Thomas S. Robertson and Harold H. Kassirjian, Englewood Cliffs, NJ: Prentice-Hall, 1-49.
- Anderson, John R. (1983), "A Spreading Activation Theory of Memory," *Journal of Verbal Learning and Verbal Behavior*, 22 (June), 261-295.
- (1993), *Rules of the Mind*, Hillsdale, NJ: Erlbaum.
- and Gordon H. Bower (1973), *Human Associative Memory*, New York: Halstead.
- and Christian Lebiere (1998), *The Atomic Components of Thought*, Mahwah, NJ: Erlbaum.
- Baeyens, Frank, Geert Crombez, Jan De Houwer, and Paul Eelen (1996), "No Evidence for Modulation of Evaluative Flavor-Flavor Associations in Humans," *Learning and Motivation*, 27 (May), 200-241.
- , Hilde Hendrickx, Geert Crombez, and Dick Hermans (1998), "Neither Extended Sequential nor Simultaneous Feature Positive Training Result in Modulation of Evaluative Flavor-Flavor Conditioning in Humans," *Appetite*, 31 (October), 185-204.
- Boush, David and Barbara Loken (1991), "A Process Tracing Study of Brand Extension Evaluation," *Journal of Marketing Research*, 28 (February), 16-28.
- Broniarczyk, Susan M. and Joseph W. Alba (1994), "The Importance of the Brand in Brand Extension," *Journal of Marketing Research*, 31 (May), 214-228.
- Chapman, Gretchen B. and Steven J. Robbins (1990), "Cue Interaction in Human Contingency Judgment," *Memory and Cognition*, 18 (September), 537-545.
- Dickinson, Anthony, David Shanks, and John Evenden (1984), "Judgement of Act-Outcome Contingency: The Role of Selective Attribution," *Quarterly Journal of Experimental Psychology*, 36A (February), 29-50.
- Erickson, Michael A. and John K. Kruschke (1998), "Rules and Exemplars in Category Learning," *Journal of Experimental Psychology: General*, 127 (June), 107-131.
- Farquhar, Peter H. and Paul M. Herr (1993), "The Dual Structure of Brand Associations," in *Brand Equity and Advertising: Advertising's Role in Building Strong Brands*, ed. David A. Aaker and Alexander L. Biel, Hillsdale, NJ: Erlbaum, 263-277.
- Feldman, Jack M. and John G. Lynch (1988), "Self-Generated Validity and Other Effects of Measurement of Belief, Attitude, Intention, and Behavior," *Journal of Applied Psychology*, 73 (August), 421-435.
- Gilbert, Daniel T. (1999), "What the Mind's Not," in *Dual-Process Theories in Social Psychology*, ed. Shelly Chaiken and Yaacov Trope, New York: Guilford, 3-11.
- Gluck, Mark A. and Gordon H. Bower (1988), "From Conditioning to Category Learning: An Adaptive Network Model," *Journal of Experimental Psychology: General*, 117 (September), 227-247.
- Greenwald, Anthony G. and Clark Leavitt (1984), "Audience Involvement in Advertising: Four Levels," *Journal of Consumer Research*, 11 (June), 581-592.
- Hutchinson, J. Wesley and Joseph W. Alba (1991), "Ignoring Irrelevant Information: Situational Determinants of Consumer Learning," *Journal of Consumer Research*, 18 (December), 325-345.
- Janiszewski, Chris and Stijn M. J. van Osselaer (2000), "A Connectionist Model of Brand-Quality Associations," *Journal of Marketing Research*, 37 (August), 331-350.
- John, Deborah Roedder, Barbara Loken, and Christopher Joiner (1998), "The Negative Impact of Extensions: Can Flagship Products Be Diluted?" *Journal of Marketing*, 62 (January), 19-32.
- Kamin, Leon J. (1969), "Predictability, Surprise, Attention, and Conditioning," in *Punishment and Aversive Behavior*, ed. Byron A. Campbell and Russell M. Church, New York: Appleton-Century-Crofts, 279-296.
- Keller, Kevin Lane (1993), "Conceptualizing, Measuring, and Managing Customer-Based Brand Equity," *Journal of Marketing*, 57 (January), 1-22.
- (1998), "Brand Knowledge Structures," in *Strategic Brand Management: Building, Measuring, and Managing Brand Equity*, Upper Saddle River, NJ: Prentice-Hall, 86-129.
- Kruschke, John K. (1992), "ALCOVE: An Exemplar-Based Connectionist Model of Category Learning," *Psychological Review*, 99 (January), 22-44.
- (1996), "Base Rates in Category Learning," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22 (January), 3-26.
- and Mark K. Johansen (1999), "A Model of Probabilistic Category Learning," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25 (September), 1083-1119.

- Loken, Barbara and Deborah Roedder John (1993), "Diluting Brand Beliefs: When Do Brand Extensions Have a Negative Impact?" *Journal of Marketing*, 57 (July), 71–84.
- Lopez, Francisco J., David R. Shanks, Julian Almaraz, and Pablo Fernandez (1998), "Effects of Trial Order on Contingency Judgments: A Comparison of Associative and Probabilistic Contrast Accounts," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24 (May), 672–694.
- McClelland, James L. (2000), "Connectionist Models of Memory," in *The Oxford Handbook of Memory*, ed. Endel Tulving and Fergus I. M. Craik, Oxford: Oxford University Press, 583–596.
- and David E. Rumelhart (1986), "A Distributed Model of Human Learning and Memory," in Rumelhart et al., eds., 170–215.
- Meyer, Robert J. (1987), "The Learning of Multiattribute Judgment Policies," *Journal of Consumer Research*, 14 (September), 155–173.
- Nedungadi, Prakash (1990), "Recall and Consumer Consideration Sets: Influencing Choice without Altering Brand Evaluations," *Journal of Consumer Research*, 17 (December), 263–276.
- Nosofsky, Robert M., Thomas J. Palmeri, and Stephen C. McKimley (1994), "Rule-Plus-Exception Model of Classification Learning," *Psychological Review*, 101 (January), 53–79.
- Park, C. Whan, Sung Youl Jun, and Allan D. Shocker (1996), "Composite Branding Alliances: An Investigation of Extension and Feedback Effects," *Journal of Marketing Research*, 33 (November), 453–466.
- , Sandra Milberg, and Robert Lawson (1991), "Evaluation of Brand Extensions: The Role of Product Feature Similarity and Brand Concept Consistency," *Journal of Consumer Research*, 18 (September), 185–193.
- Pearce, John M. (1994), "Similarity and Discrimination: A Selective Review and a Connectionist Model," *Psychological Review*, 101 (October), 587–607.
- Pham, Michel Tuan and Gita Venkataramani Johar (1997), "Contingent Processes of Source Identification," *Journal of Consumer Research*, 24 (December), 249–265.
- Price, Paul C. and J. Frank Yates (1995), "Associative and Rule-Based Accounts of Cue Interaction in Contingency Judgment," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21 (November), 1639–1655.
- Rescorla, Robert A. and Allan R. Wagner (1972), "A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement," in *Classical Conditioning II: Current Research and Theory*, ed. Abraham H. Black and William F. Prokasy, New York: Appleton-Century-Crofts, 64–99.
- Rumelhart, David E., James L. McClelland, and the PDP Research Group, eds. (1986), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, *Foundations*, Cambridge, MA: MIT Press.
- Sanbonmatsu, David M., Steven S. Posavac, Frank R. Kardes, and Susan P. Mantel (1998), "Selective Hypothesis Testing," *Psychonomic Bulletin and Review*, 5 (June), 197–220.
- Shanks, David R. (1991), "Categorization by a Connectionist Network," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17 (May), 433–443.
- , David Charles, Richard J. Darby, and Azwiyati Azmi (1998), "Configural Processes in Human Associative Learning," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24 (November), 1353–1378.
- , David Charles, Richard J. Darby, and Anthony Dickinson (1996), "Distinguishing Associative and Probabilistic Contrast Theories of Human Contingency Judgment," in *The Psychology of Learning and Motivation*, Vol. 34, *Causal Learning*, ed. David R. Shanks et al., San Diego, CA: Academic Press, 265–311.
- and Richard J. Darby (1998), "Feature and Rule-Based Generalization in Human Associative Learning," *Journal of Experimental Psychology: Animal Behavior Processes*, 24 (October), 405–415.
- , Richard J. Darby, and David Charles (1998), "Resistance to Interference in Human Associative Learning: Evidence of Configural Processing," *Journal of Experimental Psychology: Animal Behavior Processes*, 24 (April), 136–150.
- and Francisco J. Lopez (1996), "Causal Order Does Not Affect Cue Selection in Human Associative Learning," *Memory and Cognition*, 24 (July), 511–522.
- Simonin, Bernard L. and Julie A. Ruth (1998), "Is a Company Known for the Company It Keeps? Assessing the Spillover Effects of Brand Alliances on Consumer Brand Attitudes," *Journal of Marketing Research*, 35 (February), 30–42.
- Smith, Eliot R. (1996), "What Do Connectionism and Social Psychology Offer Each Other?" *Journal of Personality and Social Psychology*, 70 (May), 893–912.
- and Jamie DeCoster (1998), "Knowledge Acquisition, Accessibility, and Use in Person Perception and Stereotyping: Simulation with a Recurrent Connectionist Network," *Journal of Personality and Social Psychology*, 74 (January), 21–35.
- and Jamie DeCoster (1999), "Associative and Rule-Based Processing: A Connectionist Interpretation of Dual-Process Models," in *Dual-Process Theories in Social Psychology*, ed. Shelly Chaiken and Yaacov Trope, New York: Guilford, 323–336.
- Sujan, Mita (1985), "Consumer Knowledge: Effects on Evaluation Strategies Mediating Consumer Judgments," *Journal of Consumer Research*, 12 (June), 31–46.
- Van Hamme, Linda J., Shu-Fang Kao, and Edward A. Wasserman (1993), "Judging Interevent Relations: From Cause to Effect and from Effect to Cause," *Memory and Cognition*, 21 (November), 802–808.
- van Osselaer, Stijn M. J. and Joseph W. Alba (2000), "Consumer Learning and Brand Equity," *Journal of Consumer Research*, 27 (June), 1–16.
- Waldmann, Michael R. and Keith J. Holyoak (1992), "Predictive and Diagnostic Learning within Causal Models: Asymmetries in Cue Competition," *Journal of Experimental Psychology: General*, 121 (June), 222–236.
- Williams, Douglas A., Kerry E. Sagness, and Janice E. McPhee (1994), "Configural and Elemental Strategies in Predictive Learning," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20 (May), 694–709.