INSEAD

The Business School for the World

Social Science
Research Centre

Faculty & Research Working Paper

CAM: A Spreading Activation Network Model of Subcategory Positioning when Categorization Uncertainty is High

Joseph LAJOS
Zsold KATONA
Amitava CHATTOPADHYAY
Miklos SARVARY
2008/23/MKT/ISSRC (revised version of 2007/18/MKT)

CAM: A Spreading Activation Network Model of Subcategory Positioning when Categorization Uncertainty is High

JOSEPH LAJOS*

ZSOLT KATONA**

AMITAVA CHATTOPADHYAY***

MIKLOS SARVARY****

24 March 2008

The first two authors contributed equally, and the order of their names was determined by a coin toss. The authors thank Jim Bettman, Pierre Chandon, Jill Klein, Gilles Laurent, Art Markman, the Editor, the Area Editor, and three anonymous reviewers for helpful comments on an earlier draft, and Cécile Adam for assistance with the data collection. The authors gratefully acknowledge the financial support of R&D INSEAD and the Sasakawa Young Leaders Fellowship Fund. This paper won the Association for Consumer Research's Best Working Paper Award in 2006.

* PhD Student, Marketing, at INSEAD, Boulevard de Constance, 77305 Fontainebleau Cedex, zsolt.katona@insead.edu

** PhD Student, Marketing, at INSEAD, Boulevard de Constance, 77305 Fontainebleau Cedex, joseph.lajos@insead.edu

*** The L'Oréal Chaired Professor of Marketing - Innovation and Creativity, Professor of Marketing at INSEAD, 1 Ayer Rajah Avenue, Singapore 138676, amitava.chattopadhyay@insead.edu

**** Associate Professor of Marketing, Director, The International Centre for Learning Innovation (ICLI); Coordinator, Marketing Area, at INSEAD, Boulevard de Constance, 77305 Fontainebleau Cedex, miklos.sarvary@insead.edu

ABSTRACT

Many new products (e.g., PDA phones) share features with multiple categories, but are also substantially different from each of these categories. When consumers encounter such a product, they may create a new subcategory (e.g., smart phones) to accommodate it. In such situations, consumers must decide where within the category structure to position the new subcategory (e.g., under the PDA category or under the phone category). We develop a spreading activation model that we call the Category Activation Model (CAM) to predict where within a category structure consumers are likely to position a subcategory that they have created to accommodate a new, hybrid product. Based on this model, we hypothesize that the probability that an individual will position a new category subordinate to a particular category i is proportional to the relative number of categories that are already subordinate to i. We report the results of three studies that support this hypothesis, and provide evidence that accessibility is an underlying mechanism.

In this era of digital convergence, consumers often encounter new hybrid products that do not fit unambiguously into their existing mental categories (Gregan-Paxton, Hoeffler, and Zhao 2005; Moreau, Lehmann, and Markman 2001a; Moreau, Markman, and Lehmann 2001b). For example, when consumers encountered the Motorola Envoy, which would later be acknowledged as the first personal digital assistant (PDA), they had difficulty categorizing it, since it shared features with portable computers and organizers, yet was distinctly different from products in both of these categories (Keller, Sternthal, and Tybout 2002).

When consumers encounter a new product, such as the Envoy, which has features that differ significantly from existing categories, they respond by creating a new subcategory (e.g., the PDA subcategory) within their existing category structure for products (Sujan and Bettman 1989). If the new product could conceivably be placed into more than one existing category due to similarities (e.g., shared features) with several existing categories (i.e., categorization uncertainty is high; Gregan-Paxton et al. 2005; Moreau et al. 2001b), consumers must determine where to position the new subcategory. For example, upon encountering the Motorola Envoy, which shared features with both portable computers and personal organizers, consumers had to decide under which of these two categories to position the new PDA subcategory. Extant consumer research has not examined how this decision is made.

Understanding how consumers determine where to position a new subcategory when categorization uncertainty is high is important, since new products, including really new products, are often derived from existing product categories (Goldenberg, Mazursky, and Solomon 1999; Moreau et al. 2001a; Ward 1995). Recent examples include LG's five megapixel camera phone, which also functions as an organizer; Sony-Ericsson's Walkman phone, which also functions as an mp3 player, camera, pedometer, and organizer; and Apple's iPhone, which combines features from a wide range of different categories.

Furthermore, consumers' categorization decisions can have striking implications for firms. Research in psychology has shown that individuals use categories as bases for inductive inference and prediction (Markman 1989; Ross 1997; Smith and Medin 1981; Yamauchi and Markman 2000). In the marketing domain, consumers use information contained in existing product categories to make inferences about the features, functions, and performance of new products, and to predict the retail store departments in which they will be stocked (Gregan-Paxton et al. 2005; Gregan-Paxton and Moreau 2003). Furthermore, the manner in which consumers categorize products affects their thoughts about, attitudes toward, and overall evaluations of these products (Moreau et al. 2001b; Sujan 1985; Sujan and Bettman 1989), their memory for product features (Gregan-Paxton and Moreau 2003; Sujan and Bettman 1989), and their likelihood of recalling and subsequently choosing the products in memory-based choice (Nedungadi 1990; Nedungadi, Chattopadhyay, and Muthukrishnan 2001).

Since consumers' categorization decisions can have such profound effects, it follows that firms might benefit by influencing the manner in which consumers position subcategories for new products. If Motorola had made a concerted effort to influence the positioning of the PDA subcategory, perhaps the Envoy would have been more successful (Keller et al. 2002). In order to exert such influence, firms need to understand how consumers are naturally inclined to position the new subcategory. Such understanding could help the firm estimate the level of marketing effort needed to reinforce or alter the positioning of the subcategory to its advantage.

In this paper, we address the academically and managerially important question of how consumers position subcategories for new products when categorization uncertainty is high. We develop and test an analytical model of the process of subcategory construction under high categorization uncertainty. Our research provides insight into the process of subcategory construction in general, and the categorization of new products in particular.

CONSUMER CATEGORIZATION OF PRODUCTS

Consumer research on the categorization of products is grounded in a long literature within psychology, which shows that the use of category structures to differentiate objects is a fundamental cognitive activity (Medin and Smith 1984; Mervis and Rosch 1981; Murphy and Medin 1985; Rosch 1978; Smith and Medin 1981; see Barsalou 1992 for a review). According to this research, individuals create and maintain taxonomic representations of objects that they encounter. The goal of categorization is to create a categorical structure that maximizes the similarity of objects within each category, while simultaneously minimizing the similarity of objects in different categories (Medin and Schaffer 1978; Rosch 1978; Rosch and Mervis 1975). These categorical structures enable individuals to efficiently store and recall information learned from experience, make inferences about newly encountered objects (e.g., products), and form evaluations of them (Alba and Hutchinson 1987; Cohen 1982; Cohen and Basu 1987; Fiske 1982; Fiske and Pavelchak 1986; Gregan-Paxton and Moreau 2003; Hamilton and Sherman 1996; Sujan 1985; Sujan and Dekleva 1987). Individuals can also form categorical organizations in response to goals (e.g., “things to take out of your house in case of a fire”; Barsalou 1983, 1985, 1991). These goal-derived, or *ad hoc*, categories exist in harmony with taxonomic category structures (Coley, Medin, and Atran 1997; Medin et al. 1997; Ross and Murphy 1999).

Subtyping

Given that the goal of categorization is to classify objects within category structures that maximize within-category similarity while simultaneously minimizing between-category similarity, the following question arises: How might consumers respond to product information that suggests both significant similarities and differences with the most closely related product category? To answer this question, consumer researchers have utilized Taylor's (1981) subtyping model, which suggests that individuals respond to discrepant information about an object by constructing subcategories within more general categories when inconsistencies are large and cannot be filtered out (O'Sullivan and Durso 1984; Taylor and Crocker 1981; Weber and Crocker 1983). Consumer research has observed that consumers who are experts within a particular product category are more likely than novices to partition that product category using subcategories (Alba and Chattopadhyay 1985; Sujan 1985). Moreover, consumers' category structures for products can have several levels of categories and subcategories, with each deeper level containing a narrower, more well-defined group of products (Meyers-Levy and Tybout 1989; Sujan and Dekleva 1987).

Sujan and Bettman (1989) have conducted the most detailed investigation of consumer subcategory construction to date. In studies involving fictitious new brands of cameras, they observed that, when participants perceived a brand as being strongly discrepant from the camera category, they constructed a niche subcategory for the brand, whereas when they perceived a brand as being only moderately discrepant, they placed it directly into the camera category. Importantly, Sujan and Bettman (1989) showed that consumers construct a new subcategory in response to even just one highly discrepant brand.

Multiple Category Inferences

Although research on subtyping has generally assumed that the category into which an object will be placed is known (e.g., the camera category in Sujan and Bettman's 1989 studies), some objects are congruent with more than one category (Murphy and Ross 1994). For example, when consumers encountered the Motorola Envoy, they could have interpreted it as either a portable computer or a personal organizer since it shared features with both of these categories (Keller et al. 2002).

Only relatively recently has research examined the mental processes of individuals when they encounter objects that could potentially belong to multiple categories (Gregan-Paxton et al. 2005; Malt, Ross, and Murphy 1995; Moreau et al. 2001b; Murphy and Ross 1994, 1999; Ross and Murphy 1996). This research has focused on how the multiple categories to which the object could potentially belong each contribute to individuals' expectations and inferences about the object and their evaluations of it. For example, in the marketing literature, Gregan-Paxton et al. (2005) examined the nature of consumers' inferences, under different contingencies, when faced with a product that could belong to multiple categories (e.g., a PDA phone), and Moreau et al. (2001b) examined the formation of expectations and preferences for such products.

Goals of the Present Research

Although categorization research has established that subtyping is an important phenomenon, and more recently has begun to explore the mental processes underlying multiple category inferences, extant research has not yet connected these two important topics in a coherent way. The goal of the present research is to fill that gap by developing an analytical model of how subtyping occurs when the target object is similar to, yet substantially different from, more than one potential parent category (i.e., when categorization uncertainty is high).

Upon encountering a product that shares features with multiple categories but is also highly discrepant from each of them, Sujan and Bettman's (1989) research, discussed above, suggests that consumers will form a new subcategory, due to the significantly discrepant features. Extant research relies on similarity to predict where within a category structure people are likely to position an object, in this case a new subcategory (Barsalou 1992; Rosch 1978; Smith and Medin 1981). However, in the context of our research, similarity is not diagnostic (i.e., it does not yield a unique result), since the object being categorized shares similarities with multiple categories. Thus, going back to the example of the Motorola Envoy, which shared features both with portable computers and with personal organizers, similarity would not be diagnostic for predicting under which of these categories consumers would position the new PDA subcategory.

In this paper we develop and test an analytical process model, the Category Activation Model (CAM), to predict where within a taxonomic category structure consumers will position a subcategory for a new product when categorization uncertainty is high. In doing so, we bring together research on subtyping and multiple category inferences, advance theory within both of these domains, and demonstrate the usefulness of this new understanding in a consumer context that is highly relevant to managers who are grappling with digital convergence. Our work is particularly significant since new products are often derived from existing product categories (Goldenberg et al. 1999; Moreau et al. 2001a; Ward 1995) and yet to date there has been no research to address the issue that we discuss.

A MODEL OF SUBCATEGORY POSITIONING

Models of Categorization

Models of categorization focus on how individuals organize and represent knowledge, and the implications of these representations for inductive inferences (Medin and Schaffer 1978; Nosofsky 1986; Nosofsky and Palmeri 1997). They seek to explain how individuals will categorize a newly encountered object, and then how they will use their knowledge of this category to draw inferences about the object. These models draw on research on the learning and memory of classifications (Shepard 1964; Shepard and Chang 1963; Shepard, Hovland, and Jenkins 1961; see also Rehder and Hoffman 2005), and are based on the notion of selective attention, which proposes that in the process of classification individuals tend to narrow their attention to only those dimensions that are relevant for the given classification (Shepard et al. 1961). Most major categorization models, including exemplar models (Medin and Schaffer 1978; Nosofsky 1986; Nosofsky and Palmeri 1997), prototype models (Minda and Smith 2000; Smith and Minda 1998), and rule-based models (Smith, Patalano, and Jonides 1998), include selective attention as a key mechanism.

Category Activation Model

The CAM differs from extant models of categorization in two important ways. First, whereas extant models primarily seek to describe how category knowledge is represented and to predict into which categories objects will be placed, the goal of the CAM is to predict where entire new subcategories will be positioned when categorization uncertainty is high (i.e., when the new subcategory is created to classify an object that shares significant similarities with several categories but also differs significantly from each of them). Second, whereas extant models are largely based on the notion of selective attention, the CAM relies on category accessibility, a point that we elaborate on next.

According to models of classification and comprehension (Wyer and Radvansky 1999; Wyer and Srull 1989), when an individual encounters an object, he or she first interprets it in terms of stored knowledge, without regard to any specific processing goal. The specific representation that the individual retrieves from "permanent storage" at this early processing stage is a function of "the frequency and recency with which it has been used in the past and the degree to which its contents match" the encountered object (Wyer and Radvansky 1999, p. 92). Thus, if the encountered object (or subcategory in the present case) is ambiguous (i.e., possesses features that are congruent with several alternative categories), the individual will be more likely to position it under those categories that have been more frequently and recently used (i.e., those that are more accessible). Psychological research on the classification of ambiguous entities strongly supports this contention (Herr 1986, 1989; Higgins, Bargh, and Lombardi 1985; Higgins and King 1981; Higgins, Rholes and Jones 1977; Srull and Wyer 1979; Wyer 2007; Wyer and Srull 1981).

In this article, we examine the classification of a new subcategory for an ambiguous object that shares both similarities and differences with several categories. When a new subcategory could potentially be positioned under more than one category (because the encountered object possesses features that are congruent with more than one category), the accessibility of the competing categories becomes crucial. Thus, the CAM focuses on a context in which accessibility plays a critical role in determining the category to which the new subcategory is assigned. It is important to note that we do not suggest that similarity per se is unimportant in subcategory positioning, but rather that, given the context that we study, accessibility becomes the key arbiter of subcategory positioning.

Drawing on recent mathematical research on the growth of networks (Barabasi and Albert 1999, 2002), we next develop our model of where within an existing taxonomic category structure individuals are likely to position a new subcategory when category uncertainty is high.

Spreading Activation

We build on previous research that shows that priming a category increases its accessibility and subsequent use (Herr 1986; 1989; Higgins, Bargh and Lombardi 1985; Higgins and King 1981; Higgins, Rholes and Jones 1977; Srull and Wyer 1979; Wyer 2007; Wyer and Radvansky 1999; Wyer and Srull 1981, 1989). We assume that when a category is accessed (e.g., by thinking about a product in the category), some of the resulting activation remains with the category, and the rest spreads through the entire network. Furthermore, we assume that the increase in activation that results when a category is accessed increases the probability that a new subcategory will be positioned under it. Combining these two assumptions, it follows that if we accurately describe the process by which activation spreads through the network, we should be able to predict the probability that a new subcategory will be positioned at any relevant location within the category structure. We expect that these probabilities will depend on the existing link structure, since these links determine how activation spreads through the network.

We describe the spreading activation process as follows. Let 1, 2, …, n denote the nodes (i.e., categories) within the existing network (i.e., category structure). A link connects two nodes if a subordinate relationship exists between them (i.e., one is a subcategory of the other). We assume that the access of nodes initiates activation, which then spreads through the network. We model each time period during which no new nodes are created. For such a time period, we derive the limit activation level (i.e., the activation level at the end of the time period) of each node as the result of a discrete iterative process. In this process, node i starts with an arbitrary activation level, $a_i^{(0)}$ and $A^{(0)} = \sum_{i=1}^{n} a_i^{(0)}$ denotes the total activation in the network. At the beginning of each step in the iterative process, we assume that the activation level of each node increases by β_1 due to external activation that is uniformly distributed across the network. This assumption is equivalent to saying that, during a given time period, each existing category is accessed the same number of times on average. We use this assumption to simplify the exposition of the model,[1] but we show in appendix A that relaxing this assumption by allowing each node to receive an independent amount of external activation leads to a qualitatively identical result (i.e., the same hypothesis).

[1] In study 1, we control for the implementation of this assumption.

We assume that β_2 units of this activation remain at node i, and $a_i^{(t)} + \beta_1 - \beta_2$ units of activation spread to node i's neighbors (i.e., its subcategories and the parent category to which it is connected). To again ease exposition, we assume that these units of spreading activation are distributed equally among node i's neighbors. In the generalized model that we develop in appendix A, we show that the results are qualitatively the same when this assumption is relaxed. Finally, at the end of the iterative step, we assume that the activation level of each node decreases by β_1 such that the overall activation level of the entire network remains constant until a new subcategory is created.[2] The iterative process is represented by the following equation:

$$a_i^{(t+1)} = \beta + \frac{a_{i1}^{(t)} - \beta}{d_{i1}} + \frac{a_{i2}^{(t)} - \beta}{d_{i2}} + \ldots + \frac{a_{ik}^{(t)} - \beta}{d_{ik}}, \qquad \textbf{(1)}$$

where $\beta = \beta_2 - \beta_1$ and $i1$, $i2$, …, ik are the neighbors of node i, and d_{jk} is the degree (number of neighbors) of neighboring node j. In appendix B, we show that the series $a_i^{(t)}$ converges as $t \to \infty$ for arbitrary starting values. Furthermore, we show that convergence occurs quickly if, instead of arbitrary starting values, we use the final activation levels of all existing nodes as new starting values after a new node is created.

To complete the description of the category positioning process, we must define a new starting activation for every newly created node. If we arbitrarily assume that the total activation of the network is constant over time, then we define this starting activation level to be 0. However, this assumption is not necessary, since the proof in appendix B shows that the process converges from any starting values. Thus, we may instead assume that the activation level of the network increases or decreases as it grows. To model increasing total activation, we define the starting activation level of the new node to be positive, and all of the results

[2] This uniform decrease of activation means that 'forgetting' of the categories in the existing network is, on average, identical for each category. We relax this uniformity assumption in the generalized model (see appendix A).

still follow. We may also assume that the total activation level of the entire network decreases over time, and again the results remain unchanged.

Knowing that the iterative process in equation 1 converges, it follows that the limit activation level of any node *i,* denoted by a_i, must satisfy equation 2.

$$a_i = \beta + \frac{a_{i1} - \beta}{d_{i1}} + \frac{a_{i2} - \beta}{d_{i2}} + \ldots + \frac{a_{ik} - \beta}{d_{ik}}. \tag{2}$$

In appendix B, we show that normalizing $A^{(0)} / (2n - 2 + n\beta)$ to 1 and solving equation 2 for each node in the network yields the following limit activation level just before a new node is created:

$$a_i = d_i + \beta. \tag{3}$$

Thus, we propose that before a new node is created, the activation level of each node in the existing network is a linear function of its degree (i.e., the number of other nodes to which it is directly linked). Since, in a taxonomic category structure, the degree of each category is equal to the number of subcategories connected to it plus one (for its parent category), the CAM predicts that the activation level of each category is proportional to the number of subcategories that are connected to it.

Following a large body of previous categorization research both in marketing (Alba and Hutchinson 1987; Cohen and Basu 1987; Gregan-Paxton and Moreau 2003; Sujan 1985; Sujan and Bettman 1989; Sujan and Dekleva 1987) and psychology (Medin and Smith 1984; Mervis and Rosch 1981; Murphy and Medin 1985; Rosch 1978; Smith and Medin 1981) we have thus far assumed the use of taxonomic category structures. Nevertheless, the model can be generalized to more complex structures in which an object can be placed in multiple subcategories, subcategories can be connected to multiple parent categories, and horizontal and vertical connections can exist between categories. All of our previous results hold for this greatly generalized CAM, as shown in appendix A.

Hypothesis Development

Combining the prediction that the activation level of each category is proportional to the number of subcategories that are connected to it with our assumption that the probability that a new subcategory will be positioned under an existing category is directly related to the existing category's activation level (when the new subcategory contains an ambiguous object that shares similarities with several potential parent categories), leads to our hypothesis:

H: The probability that an individual will position a new category subordinate to a particular category *i* is proportional to the relative number of categories that are already subordinate to *i*.

What link structures should result from the process of subcategory positioning that we have proposed? As the process progresses, we are likely to observe a network in which the vast majority of subcategories are connected to a relatively small number of parent categories. The tendency for links to form between new nodes and those existing nodes that have high degrees is called *preferential attachment* by network theorists, and is a common property of many naturally occurring networks. Popular examples include the human nervous system, the worldwide web, terrorist networks, and the network of reactions in protein synthesis (Barabasi and Albert 1999).

We next report the results of three studies. In the pilot study, we analyze taxonomic category structures that lab participants generated for 100 products to test whether the overall category structures created by participants have properties congruent with our hypothesis. In study 1, we experimentally test our hypothesis in a new product context, and provide evidence in support of our assumption that accessibility is an underlying mechanism. In study 2, we differentiate between the impact of subcategory numerosity and perceived category breadth in order to rule out an alternative explanation.

PILOT STUDY: CATEGORIZATION OF 100 PRODUCTS

The purpose of the pilot study was to perform a preliminary test of our hypothesis by asking participants to build category structures for 100 products and then examining some properties of the structures that they formed. Specifically, we first test the null hypothesis that participants positioned subcategories by choosing parent categories uniformly, independent of the number of subcategories that were already connected to them. Having thus established that subcategory numerosity matters, we then test whether participants were more likely to position new subcategories under those categories to which relatively more subcategories were already connected. Finally, we test whether participants who spent less time on the task relied relatively more on the heuristic represented by the CAM as we would expect, and whether the order in which the products were presented affected the category structures that participants formed as we would expect if participants were not merely replicating a known category structure from memory.

Method

Participants. One hundred and sixteen students from a large, urban university participated in the study in exchange for a voucher for a sandwich and a soda that was redeemable at a nearby café.

Materials. We wrote a computer program that made it possible for participants to create taxonomic category structures for a large number of products. The program divided the screen into upper and lower halves. The upper half of the screen displayed product descriptions that we downloaded from a major retail website. Each description consisted of the name of the product, a photograph of the product, the product's price, and a list of the product's key functions and features. All 100 products that we selected for this study were home electronic products and thus quite similar, making the manner in which they should be categorized non-obvious (i.e., the task was ambiguous for all 100 products even though the stimuli did not include any new, hybrid products). Furthermore, given that some participants may have been familiar with the retail website, choosing all the products from just one category ensured that knowledge of the site's organization would not provide any information about how the products should be differentiated from each other. The lower half of the screen displayed the participant's category structure, which initially consisted of only one category labeled "Products," and two buttons labeled "Construct Category" and "Place Product."

Throughout the study, participants could click the "Construct Category" button at any time to add a new subcategory to the structure. Participants could construct as many or as few subcategories as they desired, position them in whatever manner they liked (the number of nested levels was not restricted), and assign whatever names they thought were appropriate. By clicking a product subcategory and then clicking the "Place Product" button, participants could place the product that was displayed in the upper half of the screen into the selected subcategory in the lower half of the screen. The name of the product would be added to the category structure, and the next product to be placed would appear in the upper half of the screen. The program recorded each participant's actions throughout the study, including the exact time at which he or she constructed each subcategory, the names and positions that he or she assigned to these subcategories, the exact time at which he or she placed each product, and the identities of the subcategories into which he or she placed them.

Procedure. A recruiter greeted students near a university and asked them if they would be interested in participating in a study in exchange for a voucher for a sandwich and a soda. Those who agreed followed the recruiter to a behavioral lab, where they read and signed a consent form and then took part in the study individually.

The experimenter assigned each participant to a small room that contained a computer on which the product categorization program had been installed. On the first two screens, participants read instructions that explained how to use the program and stated that they would be asked to build a category structure for 100 products. The last section of the instructions explained that, before beginning this task, they would be asked to practice using the program by building a category structure for 10 products that would not appear in the main task. After reading these instructions, participants advanced to the practice task, in which they used the program to build a category structure for 10 computer peripherals that appeared in a fixed order. Note that we selected the practice products from a completely different category (computer peripherals) than the 100 products (all home electronics) that we chose for the main task. After completing the practice task, participants advanced to the main task in which they used the program to build a category structure for the 100 home electronics products, which appeared in random order. Finally, participants completed a questionnaire that contained several questions about individual factors (e.g., gender, age, academic major) and an open-ended suspicion probe. Upon completion of the questionnaire, the experimenter debriefed participants and gave them each a voucher for a sandwich and a drink.

Results and Discussion

We excluded the responses of 14 participants (12% of the sample) who either did not construct any subcategories or constructed only one. The remaining 102 participants constructed 19.76 subcategories on average, and placed an average of five products into each of these subcategories.

Subcategory Positioning. In total, participants constructed nearly 2,000 subcategories. Since each of these subcategories was positioned within a different category structure (the structure that the participant had been developing), we used two special statistical tools to test our hypothesis that the probability that an individual will position a new category subordinate to a particular category *i* is proportional to the relative number of categories that are already subordinate to *i*.

First, we performed a likelihood ratio test with the null hypothesis that participants positioned subcategories by choosing parent categories uniformly, independent of the number of subcategories that were already connected to them. In appendix C, we show that the value of the likelihood ratio statistic is 775.93, which is greater than the critical value of -194.70 (p = .001). Thus, we reject the null hypothesis, thereby ruling out the possibility that participants simply selected parent categories with equal probabilities.

Second, we tested whether participants were more likely to position new subcategories under those categories that already possessed relatively large numbers of subordinates (i.e., a relatively high degree) than under those that possessed few subordinates (i.e., a low degree). We estimated the parameters $\varphi(i)$ (i = 2, .., 10) of a model in which individuals select a k-degree category with probability:

$$\frac{\varphi(k)}{\sum_{i=1}^{n}\varphi(d_i)} \tag{4}$$

where $\varphi(1)$ is normalized to 1. The $\varphi(i)$ parameters in this model represent the importance of categories with degree i in the subcategory positioning process. In particular, the CAM predicts that $\varphi(i)=(i+\beta)/(1+\beta)$ (i.e., that $\varphi(i)$ is an increasing linear function of the degree, i).

We used maximum likelihood to estimate the $\varphi(i)$ parameters for i = 2, …, 10. We did not have enough data to estimate the $\varphi(i)$ parameters for i > 10, since approximately 90% of participants did not position more than 10 subcategories under any one category.

Therefore, we excluded observations for new subcategories that were positioned under categories to which more than 10 subcategories were already connected, leaving us with 1,402 usable data points. Calculating the likelihood functions necessary to estimate the parameters for the entire data set at once is too complex. Therefore, we randomly split the participants into three groups and estimated the $\varphi(i)$ parameters separately for each group (see table 1).

Insert table 1 about here

We then tested whether the $\varphi(i)$ parameters form an increasing function, as the CAM predicts, by regressing our estimates of $\varphi(i)$ on *i*. The coefficient estimates for all three groups are positive and highly significant: 2.21 (p = .002) for group 1, 1.87 (p < .001) for group 2, and 2.21 (p < .001) for group 3. Since the results are similar for all three groups, we can use the average of the $\varphi(i)$ parameter estimates across the groups as a global estimate (see figure 1).

Insert figure 1 about here

Having found support for our hypothesis using two statistical techniques, we next examined individual differences in the category positioning process.

Time effects. For each participant, we estimated the β parameter included in the model (equations 1, 2, 3). Recall that a lower β indicates that an individual relied relatively more on the heuristic represented by the CAM when positioning new subcategories. That is, a lower β indicates that the number of subcategories already connected to a particular category had a greater influence on the probability that the individual would position a new subcategory there as well.

We used these estimates of β to investigate the relationship between the degree to which participants relied on the heuristic represented by the CAM and the average amount of time that they spent positioning each new subcategory. We predicted that participants who spent less time positioning each new subcategory would rely on the heuristic relatively more, since they would only consider a few potential parent categories—those with the most subcategories.

To test this prediction, we estimated each participant's β parameter using maximum likelihood. Then we calculated the average amount of time that each participant spent positioning new subcategories as t = (total time spent on the task / total number of categories created). Finally, we regressed β on t. The estimated coefficient is negative and significant (-9.37E9, $p < .05$), which suggests that participants who spent less time positioning each new subcategory indeed relied on the heuristic represented by the CAM relatively more, as we predicted.

Product order. Even though we purposefully selected 100 similar products so that the manner in which they should be differentiated from each other would initially be ambiguous, it is important to verify that participants did not recreate a familiar category structure from memory. Since the program displayed the 100 products in a random order, we predicted that if participants followed the heuristic represented by the CAM rather than recreating a familiar category structure, the order in which the products were presented would have affected the manner in which participants placed them in the category structure.

We tested this prediction by examining the level in the category structure and the degree of the selected category in which participants placed each product. We regressed both of these quantities on the product's position in the actual order. This regression yielded positive and significant coefficient estimates: 0.000948 ($p = .004$) for the level within the category structure, and 0.00978 ($p < .001$) for the degree of the category. These results suggest that the order in which the products were presented affected the manner in which participants categorized them, thereby providing evidence that participants did not simply recreate a familiar category structure.

Our pilot study is not without weaknesses. For example, once participants had created a subcategory, they could not eliminate it from the structure or merge it with another subcategory, since allowing them to do so would have made the analyses that we performed too complex. Due to this simplification, the study lacked some degree of realism. Furthermore, the study may have been overly long. Thus, in its later stages, participants may have begun paying less attention to the task or may have begun 'dumping' many products into the same category. Although neither of these possibilities can explain our results, a cleaner test of our hypothesis would be helpful. Additionally, although we designed the task such that the manner in which the products should be differentiated from each other would initially be ambiguous, the study did not contain any new, hybrid products. Finally, the study did not examine whether accessibility underlies the demonstrated effect.

STUDY 1: CATEGORIZATION OF A NEW HYBRID PRODUCT

The purpose of study 1 was to formally test our hypothesis in the context of a new, hybrid product. Furthermore, we sought to test our proposition that accessibility underlies the effect. We employed a 3 (priming: entertainment vs. health vs. neutral) X 2 (subcategory numerosity: entertainment vs. health) fractional factorial design. Participants in the neutral priming condition were divided equally between the two subcategory numerosity conditions, whereas all participants in the entertainment priming condition were assigned to the health subcategory numerosity condition, and all participants in the health priming condition were assigned to the entertainment subcategory numerosity condition

Method

Participants. Ninety-six students at a large, urban university participated in the study in exchange for one or two candy bars. Participants were recruited to ostensibly participate in two separate studies, and were randomly assigned to the experimental conditions described below.

Independent Variables. We manipulated the number of subcategories belonging to two broad product categories (entertainment electronics and health electronics), and controlled the frequency with which these subcategories were accessed. The stimuli included a three-level taxonomic category structure with "consumer electronics" at the top-level, "entertainment" and "health" at the middle-level, and a set of seven subcategories at the bottom-level.

We selected the seven bottom-level subcategories—home theater, computer peripherals, medical devices, dental devices, exercise equipment, nutrition aids, and massage devices—on the basis of pretest results. Specifically, we selected these particular subcategories because pretest participants were significantly more likely to categorize them as belonging to the entertainment or health categories than to a list of several other broad categories (e.g., home, office, communication). Furthermore, pretest participants were significantly more likely to categorize home theatre and computer peripherals as entertainment electronics rather than health electronics, significantly more likely to categorize medical devices and dental devices as health electronics rather than entertainment electronics, and statistically equally likely to categorize exercise equipment, nutrition aids, and massage devices as either entertainment electronics or health electronics. These results allowed us to create two experimental conditions by manipulating the position of the three subcategories that were equally likely to be categorized as health electronics or entertainment electronics.

In the entertainment subcategory numerosity condition, five subcategories were positioned under the entertainment category and two subcategories were positioned under the health category, whereas in the health subcategory numerosity condition five subcategories were positioned under the health category and two subcategories were positioned under the entertainment category (see figure 2). The titles of the seven bottom-level subcategories were held constant across the two subcategory numerosity conditions. Based on the pretest results, both of these structures were equally congruent with participants' category beliefs.

Insert figure 2 about here

We also manipulated the accessibility of the health and entertainment categories through priming. We implemented the priming manipulation, which had three levels (entertainment, health, neutral), through a proof-reading task that was described to participants as a separate study. The experimenter instructed participants to read a 750-word news story with the goal of identifying 20 misspelled words to ostensibly assess the importance of grammatical correctness in newspapers. In the entertainment prime condition, the story was titled "Survey: E-Entertainment;" whereas in the health prime condition the story was titled "America's Health Care Crisis;" and in the neutral prime condition the story was titled "Canadian-American Relations." Furthermore, in the entertainment prime condition, the story contained 20 instances of the word "entertainment" and no instances of the word "health," whereas in the health prime condition it contained 20 instances of the word "health" and no instances of the word "entertainment," and in the neutral prime condition it did not contain any instances of the words "entertainment" or "health."

Procedure. The experimenter greeted passersby in a student union building and asked them if they would be interested in participating in two short studies in exchange for one or two candy bars, depending on their performance on a proof-reading task in the first study. Those who agreed read and signed a consent form, and then took part in the experiment individually.

In the first part of the experiment, participants completed the proof-reading (i.e., priming) task. To ensure that they read their assigned stories carefully, the experimenter promised them a second candy bar if they succeeded in identifying all of the misspelled words in the story.

After participants completed the priming task, they were asked to participate in the supposed second study. In this part of the experiment, participants first familiarized themselves with their assigned category structure, which was presented on large format paper, by placing small pieces of paper with the names of 14 familiar electronics products (see figure 3) under the seven bottom-level subcategories. We chose these products on the basis of a pretest that showed that participants reliably placed two of them into each of the seven bottom-level categories. Thus, by asking participants to categorize these 14 products, we led them to access each of the seven bottom-level subcategories twice. It follows that the number of times that they accessed the middle-level entertainment and health categories differed between the two subcategory numerosity conditions. In the entertainment subcategory numerosity condition—in which five subcategories were connected to the entertainment category and two subcategories were connected to the health category—each participant accessed the entertainment category 10 times and the health category four times. In the health subcategory numerosity condition—in which two subcategories were connected to the entertainment category and five subcategories were connected to the health category—each participant accessed the entertainment category four times and the health category 10 times.

Insert figure 3 about here

Next, the experimenter presented participants with a description of a fictitious new product called the Exercise Buddy (see figure 4). The Exercise Buddy was a hybrid product that shared characteristics with both familiar entertainment electronics (e.g., stores up to 1 Gb of MP3s) and familiar health electronics (e.g., monitors number of calories burned). The results of a pretest indicated that participants were statistically equally likely to categorize the

Exercise Buddy as an entertainment electronics or health electronics product when given a list of five general categories of electronics from which to choose (communications, entertainment, health, home, office). In the experiment, participants could categorize the Exercise Buddy either by placing a small piece of paper containing its name under one of the seven bottom-level subcategories in their assigned category structure or by constructing a new bottom-level subcategory, connected either to the entertainment category or to the health category. To construct a new subcategory, participants wrote the name of the new subcategory directly onto their assigned category structure in one of two blank spaces—one with a subordinate relationship to the entertainment category and one with a subordinate relationship to the health category (see figure 2)—and then placed the piece of paper that represented the Exercise Buddy under it. The dependent measure of primary interest was whether participants positioned a new subcategory beneath the entertainment category or beneath the health category. Finally, participants completed a questionnaire that contained questions about individual factors (e.g., gender, age, academic major) and an open-ended suspicion probe. Upon completion of the questionnaire, participants were debriefed and given the reward of one or two candy bars.

Insert figure 4 about here

Results and Discussion

Based on an independent coder's judgments of responses to the suspicion probe, we excluded five participants who might have been hypothesis aware. Four of these participants guessed that the first and second studies were related, and one suggested that the number of subcategories connected to the entertainment and health categories may have influenced her positioning of a new subcategory for the Exercise Buddy. After these exclusions, the study had 91 valid participants.

All 91 of these participants categorized the 14 familiar products by placing them into the subcategories that we expected based on pretest results (see figure 3). Thus, each participant accessed each of the seven bottom-level subcategories twice (by placing two products into it) before categorizing the Exercise Buddy.

Subcategory Numerosity. Our hypothesis states that the probability that an individual will position a new category subordinate to a particular category *i* is proportional to the relative number of categories that are already subordinate to *i*. In this experiment, we varied the relative number of subcategories that were subordinate to two categories—entertainment and health. In the entertainment subcategory numerosity condition, five subcategories were subordinate to the entertainment category and two subcategories were subordinate to the health category, whereas in the health subcategory numerosity condition two subcategories were subordinate to the entertainment category and five subcategories were subordinate to the health category (see figure 2). Thus, based on our hypothesis, we predicted that among those participants who received a neutral prime and subsequently chose to construct a new subcategory for the Exercise Buddy, five out of seven, or approximately 71%, would position this new subcategory under the category that already possessed five, rather than two, subcategories.

To test our hypothesis, we first examined the subcategories constructed by participants in the entertainment subcategory numerosity condition who had received a neutral prime. Twenty-one of the 22 valid participants in this condition chose to construct a new subcategory for the Exercise Buddy. Consistent with our prediction, significantly more of these participants positioned this new subcategory under the entertainment category (N = 16) than under the health category (N = 5), ($\chi^2(1) = 5.76$, $p < .05$, $w = .52$)[3]. Furthermore, the percentage of participants who positioned the new subcategory under the entertainment

[3]We report Cohen's (1988) index of effect size, *w*.

category rather than under the health category (76%) was statistically equal to 71% ($\chi^2(1) = .27$, $p > .50$, $w = .11$), as we predicted.

We next examined the subcategories constructed by participants in the health subcategory numerosity condition who had received a neutral prime. All 23 valid participants in this condition chose to construct a new subcategory for the Exercise Buddy. Consistent with our prediction, significantly more of these participants positioned this new subcategory under the health category ($N = 17$) than under the entertainment category ($N = 6$), ($\chi^2(1) = 5.26$, $p < .05$, $w = .48$). Furthermore, the percentage of participants who positioned the new subcategory under the health category rather than under the entertainment category (74%) was statistically equal to 71% ($\chi^2(1) = .09$, $p > .75$, $w = .06$), as we predicted.

Together, the results in both the entertainment and health subcategory numerosity conditions support our hypothesis that the probability that an individual will position a new category subordinate to a particular category i is proportional to the relative number of categories that are already subordinate to i.

Priming. In the model setup, we proposed that differences in accessibility underlie the effect of subcategory numerosity on positioning of a new subcategory. Thus, we expected that the effect of the subcategory numerosity manipulation on new subcategory positioning that we observed in the first part of this experiment was due to the higher accessibility of the entertainment category in the entertainment subcategory numerosity condition, and to the higher accessibility of the health category in the health subcategory numerosity condition.

To test this, we used a priming manipulation to increase the accessibility of the category that we predicted to be *less* accessible in each condition. Specifically, we primed a group of participants in the entertainment subcategory numerosity condition with a health theme, and primed a group of participants in the health subcategory numerosity condition with an entertainment theme. We predicted that this priming manipulation would diminish the effect of the subcategory numerosity manipulation on positioning of the new subcategory.

We first examined the subcategories constructed by participants in the entertainment subcategory numerosity condition who received a health prime. All 22 valid participants in this condition constructed a new subcategory for the Exercise Buddy. Consistent with our prediction, the number of these participants who positioned this subcategory under the entertainment category ($N = 13$) was statistically equal to the number who positioned it under the health category ($N = 9$), ($\chi^2(1) = .73$, $p > .25$, $w = .18$). Furthermore, the percentage of participants in this condition who positioned the new subcategory under the entertainment category rather than under the health category (59%) was marginally significantly different from the 76% that we observed among those participants in the entertainment subcategory numerosity condition who received a neutral prime ($\chi^2(1) = 3.45$, $p < .06$, $w = .40$).

Next, we examined the subcategories constructed by participants in the health subcategory numerosity condition who received an entertainment prime. Twenty-two of the 24 valid participants in this condition chose to construct a new subcategory for the Exercise Buddy. Consistent with our prediction, the number of these participants who positioned this subcategory under the health category ($N = 12$) was statistically equal to the number who positioned it under the entertainment category ($N = 10$), ($\chi^2(1) = .18$, $p > .50$, $w = .09$). Furthermore, the percentage of participants in this condition who positioned the new subcategory under the health category rather than under the entertainment category (55%) was significantly different from the 74% that we found among those participants in the health subcategory numerosity condition who received a neutral prime ($\chi^2(1) = 4.33$, $p < .05$, $w = .44$).

In summary, the results show that nearly all participants in all conditions created a new subcategory for the new, hybrid product. Furthermore, participants who received a neutral prime were more likely to position the new subcategory under the category (entertainment or health) to which relatively more subcategories were already connected. Finally, our priming manipulation, which increased the accessibility of the category to which fewer subcategories were connected, wiped out this preferential positioning effect. Taken together, the results support our hypothesis that individuals are more likely to position a new subcategory under a category to which relatively more subcategories are already connected, and our proposition that this effect occurs because categories with more connected subcategories are more accessible.

STUDY 2: DIFFERENTIATING THE IMPACT OF NUMEROSITY AND BREADTH

Although the results of study 1 support our hypothesis and provide support for accessibility as an explanatory mechanism, one could argue that the results were instead driven by differences in perceived category breadth or abstraction across the subcategory numerosity conditions. Specifically, one could argue that the category (entertainment or health) to which relatively more subcategories were connected was not more accessible than the category with relatively fewer subcategories, but rather that participants perceived it as being broader or more abstract. If this were the case, then it is possible that participants were more likely to position a new subcategory under the category with relatively more subcategories due to this greater perceived breadth rather than to higher accessibility. Although the results of the priming manipulation in study 1 suggest that accessibility matters, it is possible that, by increasing the accessibility of the less broad category, this manipulation simply countered an opposing effect of perceived breadth. The purpose of study 2 was to rule out perceived breadth as an alternative explanation by manipulating subcategory numerosity and the breadth of each category independently. The study had a 2 (subcategory numerosity:

entertainment vs. health) X 2 (entertainment category breadth: high vs. low) X 2 (health category breadth: high vs. low) between-subjects design.

Method

Participants. Three hundred and seventy-one students at a large, urban university participated in the study in exchange for a voucher for a sandwich and a drink that was redeemable at a nearby café.

Independent Variables. As in study 1, we manipulated the number of subcategories that were connected to two broad product categories (entertainment electronics and health electronics), and controlled the frequency with which these subcategories were accessed. Additionally, we manipulated the perceived breadth of each of these two categories. As in study 1, the stimuli included a three-level taxonomic category structure with “consumer electronics” at the top-level, “entertainment” and “health” at the middle-level, and a set of seven subcategories at the bottom-level. In the present experiment, the identity of the bottom-level subcategories varied across the entertainment and health category breadth conditions.

We selected the bottom-level subcategories on the basis of three pretests. The purpose of the first pretest was to identify three broad groups of product subcategories—one group consisting of subcategories that participants strongly associated with entertainment but not with health, one group consisting of subcategories that participants strongly associated with health but not with entertainment, and one group consisting of subcategories that participants strongly and equally associated with both entertainment and health. To accomplish this goal, we asked participants to rate the degree to which they associated each of 38 different subcategories with entertainment and with health on 9-point scales, ranging from 1 (not at all associated with entertainment / health) to 9 (very associated with entertainment / health). We presented the subcategories in four random orders, and counterbalanced the order of the entertainment and health scales between surveys.

The purpose of the second pretest was to identify groups of five subcategories that together are perceived as either very broad or very narrow. Additionally, we wanted each group of five subcategories to consist of three subcategories that participants strongly and equally associated with both entertainment and health, plus either two subcategories that they strongly associated with entertainment but not with health or two subcategories that they strongly associated with health but not with entertainment. In the instructions, we defined a broad group of subcategories as being relatively unrelated or dissimilar to each other (the more unrelated or dissimilar, the broader), and we defined a narrow group of subcategories as being relatively related or similar to each other (the more related or similar, the narrower). We then asked participants to rate the breadth of 40 groups of five subcategories, which we presented in four random orders, on scales ranging from -4 (very broad) to +4 (very narrow),

The purpose of the third pretest was to identify groups of two subcategories—which were either both strongly associated with entertainment but not with health, or both strongly associated with health but not with entertainment—that together are perceived as either very broad or very narrow. Additionally, we wanted each group of two subcategories to correspond with a group of five subcategories (minus the three subcategories that were strongly associated with both entertainment and health) from pretest 2. We presented the same instructions as in pretest 2, and then asked participants to rate the breadth of 40 groups of two subcategories, which we presented in four random orders, on scales ranging from -4 (very broad) to +4 (very narrow).

Insert table 2 about here

The results of these three pretests allowed us to manipulate subcategory numerosity and category breadth independently. Table 2 lists the subcategories that were connected to the entertainment and health categories in each of the experiment's eight cells. As in study 1, in the entertainment subcategory numerosity condition, there were five subcategories connected to the entertainment category and two subcategories connected to the health category, whereas in the health subcategory numerosity condition there were two subcategories connected to the entertainment category and five subcategories connected to the health category. Furthermore, according to the pretest results, the groups of subcategories in each of the high entertainment category breadth and high health category breadth cells were all rated as being statistically equally broad, and the groups of subcategories in each of the low entertainment category breadth and low health category breadth cells were all rated as being statistically equally narrow. Finally, we counterbalanced which category (entertainment or health) appeared on the left side of the category structure.

Procedure. A recruiter greeted passersby in the streets outside of a behavioral laboratory and asked them if they would be interested in participating in a study in exchange for a voucher for a sandwich and a drink. Those who agreed read and signed a consent form, and then took part in the experiment individually.

Participants engaged in the study using a computer program that we wrote. The program's interface was similar to that which we utilized in the pilot study, yet allowed participants to engage in a completely different set of tasks. The program divided the screen into upper and lower halves. As in the pilot study, the upper half of the screen displayed product descriptions. However, the lower half of the screen displayed the participant's assigned category structure rather than a blank space where an entirely new category structured could be created.

As in study 1, participants first familiarized themselves with their assigned category structure by categorizing14 familiar electronics products (see figure 5) under the seven bottom-level subcategories. We chose these products on the basis of a pretest that showed that participants reliably placed two of the products into each of the seven bottom-level categories. However, whereas in study 1 we only showed participants the name of each product, in this study we showed them each product's name and a brief description of its key functions and features. In the experiment, we only included those 14 products that corresponded with the seven bottom-level subcategories in the participant's assigned condition. Participants categorized each product by clicking the subcategory into which they wanted to place it, and then confirming their selection in a popup box. At the conclusion of this task, participants were given the opportunity to rearrange the products in any manner that they wanted by freely dragging and dropping the products between categories.

Insert figure 5 about here

Next, the experimenter presented participants with a description of the same fictitious new, hybrid product that we utilized in study 1—the Exercise Buddy (see figure 4). Participants could categorize the Exercise Buddy either by clicking the name of one of the seven bottom-level subcategories in their assigned category structure to place it there, or by constructing a new bottom-level subcategory, connected either to the entertainment category or to the health category. To construct a new subcategory, participants could click one of two blank spaces–one with a subordinate relationship to the entertainment category and one with a subordinate relationship to the health category—and could then confirm their selection and provide a name for the subcategory in a popup box. As in study 1, the dependent measure of primary interest was whether participants positioned a new bottom-level subcategory beneath the entertainment category or beneath the health category. Finally, participants completed a computer-based questionnaire that contained questions about individual factors (e.g., gender, age, academic major) and an open-ended suspicion probe. Once participants had completed these tasks, the experimenter debriefed them and gave them each a voucher for a sandwich and a drink.

Results and Discussion

All 371 participants categorized the 14 familiar products that were presented by placing them into the bottom-level subcategories that we expected based on pretest results (see figure 5). Thus, as in study 1, each participant accessed each of the seven bottom-level subcategories in his or her assigned category structure twice before categorizing the Exercise Buddy.

Forty of the study's 371 participants placed the Exercise Buddy into one of the existing subcategories. The other 331 participants created a new subcategory for the Exercise Buddy. In order to determine whether any of our manipulations influenced the propensity of participants to create a new subcategory, we regressed a dummy variable—which took the value of 0 if a participant placed the Exercise Buddy into one of the old subcategories and 1 if the participant created a new subcategory for the Exercise Buddy—on dummy variables for subcategory numerosity (entertainment = 0, health = 1), entertainment category breadth (low = 0, high = 1), health category breadth (low = 0, high = 1), category on left side of screen (entertainment = 0, health = 1), and a constant using binary logistic regression. The analysis revealed that the effects of all four of our manipulations were not significant (see table 3A). Thus, in the remainder of the analyses, we focus only on the responses of those participants who created a new subcategory for the Exercise Buddy.

Our hypothesis states that the probability that an individual will position a new category subordinate to a particular category *i* is proportional to the relative number of categories that are already subordinate to *i*. Furthermore, we propose that categories with more subordinates are relatively more accessible than those with fewer subordinates, and that it is this difference in accessibility, rather than a difference in perceived category breadth, that underlies the influence of subcategory numerosity on positioning of a new subcategory that we observed in study 1. In the present study, we again varied the relative number of subcategories that were connected to the two categories of interest (entertainment and health), and also independently varied the breadth of the entertainment category and the breadth of the health category. We predicted that the subcategory numerosity manipulation would influence participants' positioning of a new subcategory for the Exercise Buddy above and beyond any effect of the breadth manipulations.

To test this prediction, using only those 331 observations of participants who created a new subcategory, we regressed a dummy variable—which took the value of 0 if a participant positioned the new subcategory under the entertainment category and 1 if the participant positioned the new subcategory under the health category—on dummy variables for subcategory numerosity (entertainment = 0, health = 1), entertainment category breadth (low = 0, high = 1), health category breadth (low = 0, high = 1), category on left side of screen (entertainment = 0, health = 1), and a constant using binary logistic regression. Consistent with our prediction, this analysis revealed that the effect of subcategory numerosity was significant ($\chi^2(1) = 7.2$, $p < .01$, see table 3B). The positive sign on the parameter estimate indicates that participants were more likely to position a new subcategory under the category (entertainment or health) to which more subcategories were already connected, in support of our hypothesis. Additionally, the effect of entertainment category breadth was marginally significant ($\chi^2(1) = 3.13$, $p = .07$), and the effect of health category breadth was not significant ($\chi^2(1) = .32$, $p = .57$). These results suggest that breadth mattered to some degree. However, the key finding is that the effect of subcategory numerosity was above and beyond any effect of breadth, since both numerosity and breadth were included in the regression model.

Together, these results indicate that whereas subcategory numerosity did not influence the propensity of participants to create (vs. not create) a new subcategory for the Exercise

Buddy, subcategory numerosity did influence the propensity of participants to position a new subcategory under the entertainment category or under the health category, in a manner congruent with the results of study 1. Furthermore, by manipulating subcategory numerosity and the breadth of the entertainment and health categories independently, we showed in this study that subcategory numerosity has an effect on the positioning of a new subcategory that is above and beyond any effect that is attributable to differences in perceived category breadth. Together, the results of the priming manipulation in study 1 and the results of the present study provide strong support for our proposition that the effect of subcategory numerosity on the positioning of a new subcategory is due to differences in accessibility between categories to which more or fewer subcategories are connected.

GENERAL DISCUSSION

The tendency to classify or categorize a new object within one's existing knowledge structure is a basic and ubiquitous phenomenon. Previous consumer research has demonstrated that the ways in which consumers categorize products can have striking implications for firms (Gregan-Paxton et al. 2005; Gregan-Paxton and Moreau 2003; Moreau et al. 2001b; Nedungadi 1990; Nedungadi et al. 2001; Sujan 1985; Sujan and Bettman 1989). In particular, Sujan and Bettman (1989) demonstrated that when consumers encounter a new product that shares significant similarities with an existing category, yet differs significantly from it in other ways, they deal with the incongruity by subtyping, or constructing a new subcategory. However, whereas Sujan and Bettman assumed that the general category into which a new product will be placed is known (the camera category in their studies), other research has demonstrated that objects may be congruent with more than one category, a phenomenon referred to as categorization uncertainty (Gregan-Paxton et al. 2005; Moreau et al. 2001b; Murphy and Ross 1994). In such situations, it is not enough to predict whether or not consumers will construct a subcategory for a new product. One must also predict under which relevant category they will position this subcategory. This was the focus of our research.

In this paper, we developed the Category Activation Model (CAM) to predict where within a category structure consumers will position a subcategory for a new product when categorization uncertainty is high (i.e., when the new subcategory could fit within multiple potential parent categories), due to shared similarities. Building on previous research that shows that priming a category increases its accessibility and subsequent use (Herr 1986, 1989; Higgins et al. 1977, 1985; Higgins and King 1981; Srull and Wyer 1979; Wyer 2007; Wyer and Srull 1981), we assumed that when a category is accessed (e.g., by thinking about a product in the category), some of the resulting activation remains with the category, and the rest spreads through the entire network. Furthermore, we assumed that when the activation level of a category is increased due to the category being accessed, the probability that a new subcategory will be positioned under it is increased. Combining these two assumptions, we proposed that if we were to accurately describe the process by which activation spreads through the category structure, we would be able to predict the probability that a new subcategory would be positioned at any relevant location in the structure. We then described this spreading activation process, and showed that these probabilities depend on the existing subordinate relationships in the structure. Specifically, based on our model we hypothesized that the probability that an individual will position a new category subordinate to a particular category i is proportional to the relative number of categories that are already subordinate to i, consistent with previous research that has shown that such preferential attachment is a common property of many naturally occurring networks (Barabasi and Albert 1999).

We tested the CAM's predictive ability in three studies. In a pilot study, we asked participants to build their own category structures for 100 home electronics products. We analyzed the process by which participants built their category structures, and, consistent with our hypothesis, showed that participants tended to position new subcategories under those categories to which many subcategories were already connected. In study 1, we experimentally tested the CAM in the context of a new, hybrid product. Furthermore, we included a priming manipulation to test our proposition that accessibility is an underlying mechanism. Consistent with our hypothesis, we found that the relative number of subcategories connected to two parent categories significantly influenced participants' positioning of a subcategory for a new, hybrid product within these categories. Furthermore, consistent with our proposition regarding accessibility, we showed that the priming manipulation attenuated this effect. Finally, in study 2 we ruled out an alternative explanation of our study 1 results by employing a similar procedure but manipulating subcategory numerosity and category breadth independently. The results of this study showed that subcategory numerosity affects the positioning of new subcategories above and beyond any effect that can be attributed to breadth.

This article makes a significant contribution to categorization research in both marketing and psychology by providing a framework to describe and predict how individuals position subcategories for discrepant objects when categorization uncertainty is high, thereby bringing together research on subtyping and multiple category inferences. Furthermore, by conceptualizing category structures as spreading activation networks and providing a rigorous mathematical account of the spreading activation process, this paper contributes to research on the modeling of categorization phenomena.

Moving Beyond the Product Context

The three studies that we report in this article are about consumers' categorization of products. However, we believe that the CAM's predictive ability extends beyond the product context. To examine if this is the case, in an extension study we analyzed the developing structures of user directory trees on internet servers with more than 1,600 users and over 60,000 subdirectories.

Most computer users organize their files within directory trees. For example, a professor's directory tree might contain directories for several research projects, and subdirectories containing files for each of these projects. Computer users typically develop their directory trees over time, constructing new directories as they accumulate more files. Based on our hypothesis, we predict that if the professor's research project directory has five subdirectories and his or her course directory has three subdirectories, then he/she will position a new subdirectory within the research project directory with probability 5 / 8 and within the course directory with probability 3 / 8.

The mathematical literature refers to trees that result from a process in which nodes (directories) are constructed with probabilities that are proportional to their degrees (number of connected subdirectories) as *plane oriented recursive trees* or *ordered recursive trees* (Smythe and Mahmoud 1995). Thus, based on our hypothesis, we predict that the directory trees of computer users are plane oriented recursive. To test this hypothesis, we used the result of Mori (2002), who showed that plane oriented recursive trees have a power-law degree distribution. Thus, an aggregate measure of the distribution of subdirectories within parent directories allows us to compare the collected data with this theoretical result. In appendix D, we present our detailed statistical analysis of the directory trees. The results show that user directory trees on all three servers had a power-law degree distribution, consistent with the CAM's predictions.

Furthermore, beyond the degree distribution, we directly tested our hypothesis by analyzing the developing structures of 25 computer users' directory trees over time. Using the directory time stamps, which indicate exactly when each directory was created, it is possible to trace the development of each tree from a single directory to its structure at the time of the study. Thus, the resulting data are similar to those that we collected in the pilot study. However, we analyzed them using a different approach due to the extremely large number of subdirectories. In appendix D, we describe in detail the analyses that we performed. The results are consistent with our hypothesis and with the pattern of subcategory positioning that we observed in this article's three studies. In summary, our analysis of the directory trees of computer users suggests that the CAM may be applicable beyond the product categorization context, and therefore may describe a more general process that occurs during subcategory positioning.

Managerial Implications and Future Directions

In addition to the theoretical contributions that we have discussed, the finding that the probability that individuals will position a new category subordinate to a particular relevant category is proportional to the relative number of categories that are already subordinate to that category has significant managerial applications. For example, the CAM provides a means for firms to assess how consumers are likely to position a subcategory for a new product. A firm could then use this prediction to estimate the level of marketing effort that would be needed to alter or support the manner in which consumers position the subcategory, to the firm's advantage.

Furthermore, the CAM suggests one means by which such influence might be exerted. Since, based on our findings, consumers will be more likely to position a new subcategory within a category that already has many subcategories, it follows that advertising that prompts consumers to think about several subcategories within a particular domain should increase the likelihood that they will position the new subcategory within that domain.

Returning again to the example of the Motorola Envoy, Keller et al. (2002) explain that consumers had difficulty categorizing the world's first PDA, since it shared features with portable computers and organizers, yet was distinctly different from products in both of these categories. Keller et al. argue that consumers' indecisiveness regarding the categorization of the Envoy was a key factor leading to its ultimate failure. That is, if Motorola had made a concerted effort to influence consumers to create a new subcategory for the Envoy under *either* the portable computer *or* organizer category, the product may have been more successful. Motorola could have utilized the CAM in its marketing effort in two ways. First, if Motorola had a general notion that consumers would be likely to associate the Envoy with multiple existing categories (e.g., portable computers, organizers, pagers), it could have used the CAM to quickly assess the relative strength of each connection by asking consumers to list the subcategories that they associated with each of these parent categories. Furthermore, once Motorola had determined which connection was likely to be strongest and by what margin, it could have reinforced this connection, not only through traditional advertising, but also through ads and interactive activities that increased the accessibility of the target category (e.g., portable computers) by leading consumers to recall and elaborate on several of its subcategories. In summary, the CAM provides a means for marketers to both assess and influence the manner in which consumers are likely to position new subcategories when categorization is high, as in the case of new, hybrid products.

APPENDIX A: GENERALIZING THE CAM

In this section, we generalize the CAM to non-taxonomic category structures in which an object may be placed in multiple categories and a subcategory may be connected to multiple parent categories. Furthermore, we allow these connections to differ in strength, and we relax the assumption that external activation is distributed uniformly throughout the network

As in the taxonomic version of the CAM, in the generalized model we represent the category structure using a network in which nodes correspond to categories and links correspond to subordinate relationships between categories. In the generalized model, we account for the direction of the subordinate relationships by directing the links from each category to its subcategories. Furthermore, we assign a weight to each link to represent the strength of the subordinate relationship. Each subcategory must have at least one in-link (parent category), but may have multiple in-links. Finally, we assume that there are no directed cycles in the network. That is, we assume that a category may not be subordinate to itself. We define the generalized degree of node *i*, d_i, as the sum of the weights assigned to its links (in- and out-). Let S denote the sum of degrees across the network (i.e., the sum of weights on all the links).

We directly generalize the iterative process described by the CAM. At the beginning of each step in the process, the activation of each node *i* increases by β_1^i due to external activation. Note that by denoting the increase in activation by β_1^i rather than by β_1 we eliminate our previous assumption that external activation is distributed uniformly throughout the network. That is, in the generalized model, the amount of external activation may differ between nodes.

We then assume that β_2 units of this activation remain at node *i*, and $a_i^{(t)} + \beta_1^i - \beta_2$ units of activation spread to node *i*'s neighbors (the nodes to which it is linked). We relax our previous assumption that this spreading activation is distributed equally among node *i*'s neighbors by now assuming instead that it is distributed proportional to the weights on the links between node *i* and each of its neighbors.

Finally, we assume that the activation level of each node decreases by β_1^i due to loss of activation. With only a few minor mathematical assumptions, it can be shown that the spreading activation process converges such that normalizing $A^{(0)} / \sum_{i=1}^{n} \beta_i$ to 1, the limit activation of node *i* is:

$$a_i = d_i + \beta_i, \qquad \textbf{(5)}$$

where $\beta_i = \beta_1^i - \beta_2$. Note that β_i can vary between nodes. These differences represent differences in the external activation of each node. For example, if categories with many subcategories are accessed relatively more frequently, we may have $\beta_i = d_i$, leading to $a_i = 2d_i$, thereby accentuating the result that categories with more subordinates have a relatively higher limit activation level.

Given the limit activations in equation 5, we can predict that the probability that a new generalized subcategory will be positioned subordinate to any category *i* in the network is proportional to its generalized degree (d_i) plus its external activation coefficient (β_i). Note, however, that the model does not predict the number of parent categories to which a new generalized subcategory will be connected.

APPENDIX B: PROVING THE CONVERGENCE OF ACTIVATION LEVELS

We prove that the limit solution of the iterative process presented in equation 1 is

$$a_i = d_i + \beta, \qquad \textbf{(6)}$$

as claimed in equation 3. In the proof we do not focus on the activation level of any particular node *i*, $a_i^{(t)}$, but rather on the part of this activation that spreads across the network. The effective activation that spreads across the network is $a_i^{(t)} - \beta$, following equation 1 when written in this form:

$$a_i^{(t+1)} - \beta = \frac{a_{i1}^{(t)} - \beta}{d_{i1}} + \frac{a_{i2}^{(t)} - \beta}{d_{i2}} + \ldots + \frac{a_{ik}^{(t)} - \beta}{d_{ik}}. \quad \textbf{(7)}$$

To show convergence, we add a damping factor $\delta > 0$ to the model, thereby extending equation 7 to:

$$\delta + \left(a_i^{(t+1)} - \beta\right) = \left(\frac{a_{i1}^{(t)} - \beta}{d_{i1}} + \frac{a_{i2}^{(t)} - \beta}{d_{i2}} + \ldots + \frac{a_{ik}^{(t)} - \beta}{d_{ik}}\right)(1 - \delta). \quad \textbf{(8)}$$

The damping factor can be thought of as a natural error that occurs during the spread of activation, such that δ part of the activation does not spread over the links but rather spreads randomly throughout the network. In order to proceed with the proof, we define a weighted directed network on the basis of the undirected network that represents the category structure. For every undirected link we put a directed link in both directions with weights $1-\delta$ for each. Then we add a directed link from every node to every other node with weight δ (if there has already been a link with weight $1-\delta$, then we add δ such that the link has a weight of 1). We will use this network to represent the transition probabilities of the activation. In every step, the effective activation of a node spreads through the directed links proportionally to the weights on these links. It follows that in each step $1-\delta$ part of the activation spreads through the links defined by the category structure and δ part spreads randomly. According to Langeville and Meyer (2004), this process converges to a unique stationary distribution. Thus, if we find the stationary distribution, it must be the limit. $a_i = d_i + \beta$ is a stationary point for $\delta = 0$, since it satisfies equation 2. Thus, the limit activation as $\delta \to 0$ and $t \to \infty$ must be $a_i = d_i + \beta$. Furthermore, Langeville and Meyer (2004) show that the convergence is exponential in *t*. This suggests that convergence occurs quickly.

APPENDIX C: LIKELIHOOD-RATIO TEST FOR PILOT STUDY

In the pilot study, we used a likelihood-ratio test to reject the null hypothesis that participants positioned new subcategories uniformly under existing categories. The data that we analyzed consist of the degree of the category to which each new subcategory was connected. We denote these data points by $X_1, X_2, ..., X_n$. Furthermore, let L_0 denote the likelihood of a particular outcome given that the process is uniform according to the null hypothesis and L_1 denote the likelihood given our hypothesis that the probability that an individual will position a new category subordinate to a particular category *i* is proportional to the relative number of categories that are already subordinate to *i*. Then we use the following statistic to reject the null hypothesis:

$$S = \log(L_1 / L_0) = \sum_{i=1}^{n} \log(X_i / 2) \tag{9}$$

Note that this is not a conventional likelihood-ratio test, which would be very difficult to calculate. In order to determine the distribution of *S* under the null hypothesis, we ran simulations. These resulted in a critical value of -194.70 ($p = 0.001$) after 10,000 iterations.

APPENDIX D: STATISTICAL ANALYSIS OF SERVER DIRECTORY TREES

Directory Structure. In mathematics, degree distribution is often used to describe the qualitative properties of networks. In particular, plane oriented recursive trees are known to have a power-law degree distribution, such that:

$$L(k) \approx k^{-\gamma_1}, \tag{10}$$

where $L(k)$ denotes the proportion of nodes with degree at least k (Smythe and Mahmoud 1995). Mori (2002) showed that if the probability that a new node (node $n + 1$) will be connected to a particular existing node (node *i*) is proportional to the degree of the existing node (d_i) plus a constant (β), then:

$$\gamma_1 = 2 + \beta. \tag{11}$$

Using this result, we tested whether directory trees constructed by more than 1,600 computer users on three internet servers have a power-law degree distribution, and we estimated γ_1. We also analyzed the positioning of more than 3,000 new directories by 25 users using time stamps associated with each directory.

Data. With the consent of information systems managers, we wrote a program to collect data on the developing structure of users' directory trees on two faculty internet servers at large universities and on a student server at a high school. The program recorded a snapshot of all user directory trees on each server, and stored the creation time of each directory within each user's tree. Our data consist of 14,674 directories created by 115 users on the first university's faculty server (server 1), 17,982 directories created by 137 users on the second university's faculty server (server 2), and 35,638 directories created by 1,412 users on the high school's student server (server 3).

Degree Distribution. Our hypothesis states that the probability that an individual will position a new category subordinate to a particular category *i* is proportional to the relative number of categories that are already subordinate to *i*. Based on this hypothesis, we predicted that the directory trees of computer users have a power-law degree distribution. We tested this prediction by log-transforming both sides of equation 10, and performing OLS regression on the empirical degree distributions of the three servers. R^2 exceeded .9 for two of the regressions and exceeded .6 for the third, suggesting that user directory trees on all three servers had a power-law degree distribution, as we predicted.

Furthermore, estimates of the power function exponent (γ_1) are similar for users' directory trees on all three servers ($\hat{\gamma}_{1,1}$ = 1.62, $\hat{\gamma}_{1,2}$ = 1.13, $\hat{\gamma}_{1,3}$ = 2.20), and roughly correspond to Barabasi and Albert's (1999) finding that many naturally occurring networks have a power law degree distribution with an exponent of approximately 2. Overall, the data are consistent with our prediction that directory trees are plane oriented recursive, and with previous empirical research on random networks, thereby supporting our hypothesis.

Directory Positioning. Next, we directly tested our hypothesis by analyzing the developing structures of 25 computer users' directory trees. Using the time stamps, which indicate exactly when each directory was created, it is possible to trace the development of each tree from a single directory to its structure at the time of the study. Each time a user constructs a new directory, he or she must decide whether or not to position the new directory as a subordinate to each of the existing directories within the directory tree. According to the CAM, the user's decision for the i^{th} new directory that he or she constructs is a random variable (X_i) with probabilities that are proportional to the relative number of directories that are subordinate to each of the relevant existing directories (d_{X_i}).

For each new directory that a particular user constructed, we recorded the number of subordinate directories that previously existed within the parent directory (the degree of the parent node). Thus, for each directory construction (X_i) we captured a single realization (x_i) and its degree (d_{x_i}). We approximated the distribution of d_{X_i} by assuming that each user's actual directory tree had a power-law degree distribution ($L(k) \approx k^{-\gamma_1}$), as our previous analyses indicated. Next, we show that the relevant existing directories that the user considered as potential parent directories for a new directory also have a power-law degree distribution with an exponent of $\gamma_2 = 1 + \beta$. In other words, this exponent is exactly one less than that of the degree distribution of the entire directory tree ($\gamma_2 = \gamma_1 - 1$).

Assume that each user's actual directory tree has a power-law degree distribution ($L(k) \approx k^{-\gamma_1}$). We then calculate the probability that the degree of a randomly selected new node is greater than or equal to k (where $k \geq 1$) by summing the probabilities assigned to nodes with higher degrees:

$$P\left(d_{X_i} \geq k\right) = \sum_{d_j \geq k} \frac{d_j + \beta}{2i - 2 + i\beta} = \frac{1}{2i - 2 + i\beta} \sum_{d_j \geq k} \left(d_j + \beta\right) = \frac{1}{2i - 2 + i\beta} \left((\beta + k) L(k) + \sum_{l \geq k} L(l) \right). \quad \textbf{(12)}$$

It follows that as $i \to \infty$, $P\left(d_{X_i} \geq k\right) \sim c' k^{-\gamma_2}$, where $\gamma_2 = 1 + \beta$.

Since we examined each user's construction of all i directories in his or her directory tree, our sample for each user consists of a single realization (x_i) and its degree (d_{x_i}) for each of these i decisions (X_i). Although our sample for each user consists of observations on i different random variables, we used the pooled sample to empirically test the cumulative distribution function. We compared the degree distribution of the parent directories to the degree distribution of the entire directory tree.

If the user constructed the new directories uniformly, such that the number of subordinates of potential parent directories did not influence his or her decisions, then γ_1 should equal γ_2. However, if the probability that a user constructed a new directory as a subordinate of a particular directory was proportional to the relative number of categories that were already subordinate to it, then the difference between γ_1 and γ_2 should equal 1

We tested this prediction by estimating γ_1 and γ_2 for the 25 most developed directory structures in our sample (i.e., those with the greatest number of directories). Consistent with our prediction, γ_1 and γ_2 were significantly different for all 25 users in our sample, with a mean difference of .844. These results suggest that the pattern of new subdirectory positioning is consistent with our hypothesis, and with the pattern of subcategory positioning that we observed in the pilot study and studies and 1 and 2.

REFERENCES

Alba, Joseph W. and Amitava Chattopadhyay (1985), "Reducing the Size of the Retrieval Set: The Effects of Context and Part-Category Cues on the Recall of Competing Brands," *Journal of Marketing Research*, 22 (August), 340-349.

Alba, Joseph W. and J. Wesley Hutchinson (1987), "Dimensions of Consumer Expertise," *Journal of Consumer Research*, 13 (March), 411-454.

Barabasi, Albert-Laszlo and Reka Zsuzsanna Albert (1999), "Emergence of Scaling in Random Networks," *Science*, 286, 509-512.

____ (2002), "Statistical Mechanics of Complex Networks," *Reviews of Modern Physics*, 74 (1) 47-96.

Barsalou, Lawrence W. (1983), "Ad Hoc Categories," *Memory and Cognition*, 11, 211-227.

____ (1985), "Ideals, Central Tendency, and Frequency of Instantiation as Determinants of Graded Structure in Categories," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 629-649.

____ (1987), "The Instability of Graded Structure in Concepts," in *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*, ed. Ulric Neisser, New York, NY: Cambridge University Press, 101-140.

____ (1991), "Deriving Categories to Achieve Goals" in *The Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. 27, ed. G. H. Bower, New York, NY: Academic Press, 1-64.

____ (1992), *Cognitive Psychology: An Overview for Cognitive Scientists*. Hillsdale, NJ: Erlbaum.

Cohen, Jacob (1988), *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed., New York: Academic Press.

Cohen, Joel B. (1982), "The Role of Affect in Categorization: Towards a Reconsideration of the Concept of Attitude," in *Advances in Consumer Research*, Vol. 9, ed. Andrew A. Mitchell, Ann Arbor, MI: Association for Consumer Research, 94-100.

Cohen, Joel B. and Kunal Basu (1987), "Alternative Models of Categorization," *Journal of Consumer Research*, 13 (March), 455-472.

Coley, John D., Douglas L. Medin, and Scott Atran (1997), "Does Rank have its Privilege? Inductive Inference within Folkbiological Taxonomies," *Cognition*, 64, 73-112.

Fiske, Susan T. (1982), "Schema-Triggered Affect: Applications to Social Perception," in *Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition*, ed. Margaret S. Clark and Susan T. Fiske, Hillsdale, NJ: Erlbaum, 55-78.

Fiske, Susan T. and Mark A. Pavelchak (1986), "Category-Based Versus Piecemeal-Based Affective Responses: Developments in Schema-Triggered Affect," in *The Handbook of Motivation and Cognition: Foundations of Social Behavior*, ed. Richard M. Sorrentiono and E. Tory Higgins, New York: Guilford Press.

Goldenberg, Jacob, David Mazursky, and Sorin Solomon (1999), "Toward Identifying the Inventive Templates of New Products: A Channeled Ideation Approach," *Journal of Marketing Research*, 36 (May), 200-210.

Gregan-Paxton, Jennifer, Steve Hoeffler, and Min Zhao (2005), "When Categorization is Ambiguous: Factors that Facilitate the Use of a Multiple Category Inference Strategy," *Journal of Consumer Psychology*, 15 (2), 127-140.

Gregan-Paxton, Jennifer and Deborah Roedder John (1997), "Consumer Learning by Analogy: A Model of Internal Knowledge Transfer," *Journal of Consumer Research*, 24 (3), 266-284.

Gregan-Paxton, Jennifer and C. Page Moreau (2003), "How Do Consumers Transfer Existing Knowledge? A Comparison of Analogy and Categorization Effects," *Journal of Consumer Psychology*, 13 (4), 422-430.

Hamilton, David L. and Steven J. Sherman (1996), "Perceiving Persons and Groups," *Psychological Review*, 103 (2), 336-355.

Herr, Paul M. (1986), "Consequences of Priming: Judgment and Behavior," *Journal of Personality and Social Psychology*, 51 (6), 1106-1115.

____ (1989), "Priming Price: Prior Knowledge and Context Effects," *Journal of Consumer Research*, 16 (June), 67-75.

Higgins, E. Tory, John A. Bargh, and Wendy Lombardi (1985), "Nature of Priming Effects on Categorization," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11 (1), 59-69.

Higgins, E. Tory and Gillian A. King (1981), "Accessibility of Social Constructs: Information-Processing Consequences of Individual and Contextual Variability," in *Personality, Cognition, and Social Interaction*, ed. Nancy Cantor and John F. Kihlstrom, Hillsdale, NJ: Earlbaum, 69-121.

Higgins, E. Tory, William S. Rholes, and Carl R. Jones (1977), "Category Accessibility and Impression Formation," *Journal of Experimental Social Psychology*, 13 (March), 141-154.

Keller, Kevin Lane, Brian Sternthal, and Alice Tybout (2002), "Three Questions You Need to Ask About Your Brand," *Harvard Business Review*, 80 (9), 80-6.

Kruschke, John K. (1992), "ALCOVE: An Exemplar-Based Connectionist Model of Category Learning," *Psychological Review*, 99 (1), 22-44.

Langville, Amy N. and Carl D. Meyer (2004), "Deeper inside Page Rank", *Internet Mathematics,* 1 (3), 335-400.

Malt, Barbara C., Brian Ross, and George Murphy (1995), "Predicting Features for Members if Natural Categories when Categorization is Uncertain," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21 (3), 646-661.

Markman, Ellen M. (1989), *Categorization and Naming in Children: Problems of Induction*. Cambridge, MA: MIT Press.

Medin, Douglas L. and Stephen M. Edelson (1988), "Problem Structure and the Use of Base-Rate Information from Experience," *Journal of Experimental Psychology: General*, 117, 68-85.

Medin, Douglas L., Elizabeth B. Lynch, John D. Coley, and Scott Atran (1997), "Categorization and Reasoning Among Tree Experts: Do All Roads Lead to Rome?" *Cognitive Psychology*, 32, 49-96.

Medin, Douglas L. and Marguerite M. Schaffer (1978), "Context Theory of Classification," *Psychological Review*, 85 (May), 207-238.

Medin, Douglas L. and Edward E. Smith (1984), "Concepts and Concept Formation," *Annual Review of Psychology*, 35, 113-138.

Mervis, Carolyn B. and Eleanor Rosch (1981), "Categorization of Natural Objects," *Annual Review of Psychology*, 32, 89-115.

Meyers-Levy, Joan and Alice M. Tybout (1989), "Schema Congruity as a Basis for Product Evaluation," *Journal of Consumer Research*, 16 (June), 9-54.

Minda, John Paul and J. David Smith (2000), "Prototypes in Category Learning: The Effects of Category Size, Category Structure, and Stimulus Complexity," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 775-799.

Moreau, C. Page, Donald R. Lehmann, and Arthur B. Markman (2001a), "Entrenched Knowledge Structures and Consumer Response to New Products," *Journal of Marketing Research*, 38 (February), 14-29.

Moreau, C. Page, Arthur B. Markman, and Donald R. Lehmann (2001b), "What Is It?' Categorization Flexibility and Consumers' Responses to Really New Products," *Journal of Consumer Research,* 27 (March), 489-498.

Mori, Tamas F. (2002), "On Random Trees," *Studia Scientiarum Mathematicarum Hungarica*, 39, 143-155.

Murphy, George and Brian Ross (1994), "Predictions from Uncertain Categorizations," *Cognitive Psychology*, 27 (October), 148-193.

____ (1999), "Induction with Cross-Classified Categories," *Memory and Cognition*, 27, 1024-1041.

Murphy, Gregory L. and Douglas L. Medin (1985), "The Role of Theories in Conceptual Coherence," *Psychological Review*, 92, 289-316.

Nedungadi, Prakash (1990), "Recall and Consumer Consideration Sets: Influencing Choice without Altering Brand Evaluations," *Journal of Consumer Research*, 17 (December), 263-76.

Nedungadi, Prakash, Amitava Chattopadhyay, and A. V. Muthukrishnan (2001), "Category Structure, Brand Recall, and Choice," *International Journal of Research in Marketing*, 18 (September), 191-202.

Nosofsky, Robert M. (1986), "Attention, Similarity, and the Identification-Categorization Relationship," *Journal of Experimental Psychology: General*, 115, 39-57.

Nosofsky, Robert M. and Thomas J. Palmeri (1997), "An Exemplar-Based Random Walk Model of Speeded Classification," *Psychological Review*, 104 (2), 266-300.

Nosofsky, Robert M., Thomas J. Palmeri, and Stephen C. McKinley (1994), "Rule-Plus-Exception Model of Classification Learning," *Psychological Review*, 101 (1), 53-79.

O'Sullivan, Chris S. and Francis T. Durso (1984), "Effect of Schema-Incongruent Information on Memory for Stereotypical Attitudes" *Journal of Personality and Social Psychology*, 47 (July), 55-70.

Rehder, Bob and Aaron B. Hoffman (2005), "Eye Tracking and Selective Attention in Category Learning," *Cognitive Psychology*, 51, 1-41.

Rosch, Eleanor (1978), "Principles of Categorization," in *Cognition and Categorization*, ed. Eleanor Rosch and Barbara B. Lloyd, Hillsdale, NJ: Erlbaum.

Rosch, Eleanor and Carolyn B. Mervis (1975), "Family Resemblances: Studies in the Internal Structure of Categories," *Cognitive Psychology*, 7, 573-605.

Rosch, Eleanor, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson, and Penny Boyes-Braem (1976), "Basic Objects in Natural Categories," *Cognitive Psychology*, 8, 382-439.

Ross, Brian H. (1997), "The Use of Categories Affects Classification," *Journal of Memory and Language*, 37 (August), 240-267.

Ross, Brian H. and Gregory L. Murphy (1996), "Category-Based Predictions: Influence of Uncertainty and Feature Associations," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22 (3), 736-753.

____ (1999), "Food for Thought: Cross-Classification and Category Organization in a Complex Real-World Domain," *Cognitive Psychology*, 38, 495-553.

Shepard, Roger N. (1964), "Attention and the Metric Structure of the Stimulus Space," *Journal of Mathematical Psychology*, 1, 54-87.

Shepard, Roger N. and Jih-Jie Chang (1963), "Stimulus Generalization in the Learning of Classifications," *Journal of Experimental Psychology*, 65 (1), 94-102.

Shepard, Roger N., Carl I. Hovland, and Herbert M. Jenkins (1964), "Learning and Memorization of Classifications," *Psychological Monographs: General and Applied*, 75 (13), 1-41.

Smith, Edward E. and Douglas L. Medin (1981), *Categories and Concepts*. Cambridge, MA: Harvard University Press.

Smith, Edward.E., Andrea L. Patalano, and John Jonides (1998), "Alternative Strategies of Categorization," *Cognition*, 65, 167-196.

Smith, J. David and John Paul Minda (1998), "Prototypes in the Mist: The Early Epochs of Category Learning," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1411-1436.

Smythe, Robert T. and Hosam M. Mahmoud (1995), "A Survey of Recursive Trees," *Theoretical Probability and Mathematical Statistics*, 51, 1-27.

Srull, Thomas K. and Robert S. Wyer Jr. (1979), "The Role of Category Accessibility in the Interpretation of Information About Persons: Some Determinants and Implications," *Journal of Personality and Social Psychology*, 37 (10), 1660-1672.

Sujan, Mita (1985), "Consumer Knowledge: Effects on Evaluation Strategies Mediating Consumer Judgments," *Journal of Consumer Research*, 12 (June), 31-46.

Sujan, Mita and James R. Bettman (1989), "The Effects of Brand Positioning Strategies on Consumers' Brand and Category Perceptions: Some Insights from Schema Research," *Journal of Marketing Research*, 26 (November), 454-67.

Sujan, Mita and Christine Dekleva (1987), "Product Categorization and Inference Making: Some Implications for Comparative Advertising," *Journal of Consumer Research*, 14 (December), 372-378.

Taylor, Shelley E. (1981), "A Categorization Approach to Stereotyping," in *Cognitive Processes in Stereotyping and Intergroup Behavior,* ed. David L. Hamilton, Hillsdale, NJ: Erlbaum, 88-114.

Taylor, Shelley E. and Jennifer Crocker (1981), "Schematic Bases of Social Information Processing," in *Social Cognition: The Ontario Symposium*, Vol. 1, ed. E. Tory Higgins, C. Peter Herman, and Mark P. Zanna, Hillsdale, NJ: Erlbaum, 89-134.

Ward, Thomas B. (1995), "What's Old about New Ideas?" in *The Creative Cognition Approach*, ed. Steven M. Smith, Thomas B. Ward, and Ronald A. Fiske, Cambridge, MA: MIT Press, 157-78.

Weber, Renee and Jennifer Crocker (1983), "Cognitive Processes in the Revision of Stereotypic Beliefs," *Journal of Personality and Social Psychology,* 45 (November), 961-977.

Wyer, Robert S. Jr. (2007), "The Role of Knowledge Accessibility in Cognition and Behavior: Implications for Consumer Information Processing," in *Handbook of Consumer Psychology*, ed. Curtis Haugvedt, Frank R. Kardes, and Paul M. Herr, Mahwah, NJ: Erlbaum.

Wyer, Robert S. Jr. and Gabriel A. Radvansky (1999), "The Comprehension and Validation of Social Information," *Psychological Review*, 106 (1), 89-118.

Wyer, Robert S. Jr. and Thomas K. Srull (1981), "Category Accessibility: Some Theoretical and Empirical Issues Concerning the Processing of Social Stimulus Information," in *Social Cognition: The Ontario Symposium*, Vol. 1, ed. E. Tory Higgins, C. Peter Herman, and Mark P. Zanna, Hillsdale, NJ: Erlbaum,

____ (1989), *Memory and Cognition in its Social Context*, Hillsdale, NJ: Erlbaum.

Yamauchi, Takashi and Arthur B. Markman (2000), "Inference Using Categories," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26 (May), 776-795.

TABLE 1

MAXIMUM LIKELIHOOD ESTIMATES OF $\varphi(i)$ FOR i = 2, …, 10 FOR RANDOMLY SPLIT GROUPS 1 - 3

i	2	3	4	5	6	7	8	9	10
Group 1	1.468048	2.175498	3.281297	4.963362	5.829788	6.772659	11.39039	11.36966	27.39863
Group 2	1.771725	4.27175	6.651391	5.166867	8.092616	11.83401	14.04493	14.75348	17.65446
Group 3	1.66803	3.286901	3.720388	7.362926	9.327925	13.23313	13.35978	13.19158	21.21927

TABLE 2
STUDY 2 SUBCATEGORIES BY CONDITION

Condition			Subcategories connected to the entertainment category	Subcategories connected to the health category
High entertainment subcategory numerosity	High entertainment category breadth	High health category breadth	DVD players Ice cream makers Bicycle computers Handheld electric massagers Dive computers	Humidifiers Juicers
		Low health category breadth	DVD players Ice cream makers Bicycle computers Handheld electric massagers Dive computers	Electronic food scales Electronic body scales
	Low entertainment category breadth	High health category breadth	GPS systems Two-way radios Bicycle computers Handheld electric massagers Dive computers	Humidifiers Juicers
		Low health category breadth	GPS systems Two-way radios Bicycle computers Handheld electric massagers Dive computers	Electronic food scales Electronic body scales
High health subcategory numerosity	High entertainment category breadth	High health category breadth	DVD players Ice cream makers	Humidifiers Juicers Bicycle computers Handheld electric massagers Dive computers
		Low health category breadth	DVD players Ice cream makers	Electronic food scales Electronic body scales Bicycle computers Handheld electric massagers Dive computers
	Low entertainment category breadth	High health category breadth	GPS systems Two-way radios	Humidifiers Juicers Bicycle computers Handheld electric massagers Dive computers
		Low health category breadth	GPS systems Two-way radios	Electronic food scales Electronic body scales Bicycle computers Handheld electric massagers Dive computers

TABLE 3

A. STUDY 2 LOGISTIC REGRESSION RESULTS WITH DV = CREATED NEW SUBCATEGORY VS. NOT

Variable	*B*	Wald	Significance	exp(*B*)
Subcategory numerosity	-.094	.079	.778	.910
Entertainment breadth	.094	.079	.778	1.099
Health breadth	-.005	.000	.988	.995
Category on left	.132	.155	.694	1.141
Constant	-2.180	32.965	.000	.113

Note—$N = 371$; Dependent variable = 1 if a participant created a new subcategory for the Exercise Buddy, 0 if he/she placed the Exercise Buddy into an existing subcategory; Subcategory numerosity = 1 if the health category had more subcategories, 0 if the entertainment category had more subcategories; Entertainment breadth = 1 if breadth of the entertainment category was high, 0 otherwise; Health breadth = 1 if breadth of the health category was high, 0 otherwise; Category on left = 1 if the health category was on left side of screen, 0 otherwise.

B. STUDY 2 LOGISTIC REGRESSION RESULTS WITH DV = POSITIONED NEW SUBCATEGORY UNDER ENTERTAINMENT VS. UNDER HEALTH

Variable	*B*	Wald	Significance	exp(*B*)
Subcategory numerosity	.604	7.211	.007	1.829
Entertainment breadth	-.398	3.131	.077	.672
Health breadth	.128	.323	.570	1.136
Category on left	-.047	.044	.833	.954
Constant	.065	.066	.797	1.067

Note—$N = 331$; Dependent variable = 1 if a participant positioned a new subcategory for the Exercise Buddy under the health category, 0 if he/she positioned it under the entertainment category; Subcategory numerosity = 1 if the health category had more subcategories, 0 if the entertainment category had more subcategories; Entertainment breadth = 1 if breadth of the entertainment category was high, 0 otherwise; Health breadth = 1 if breadth of the health category was high, 0 otherwise; Category on left = 1 if the health category was on left side of screen, 0 otherwise.

FIGURE 1

INCREASING RELATIONSHIP BETWEEN $\varphi(i)$ AND DEGREE, *i*

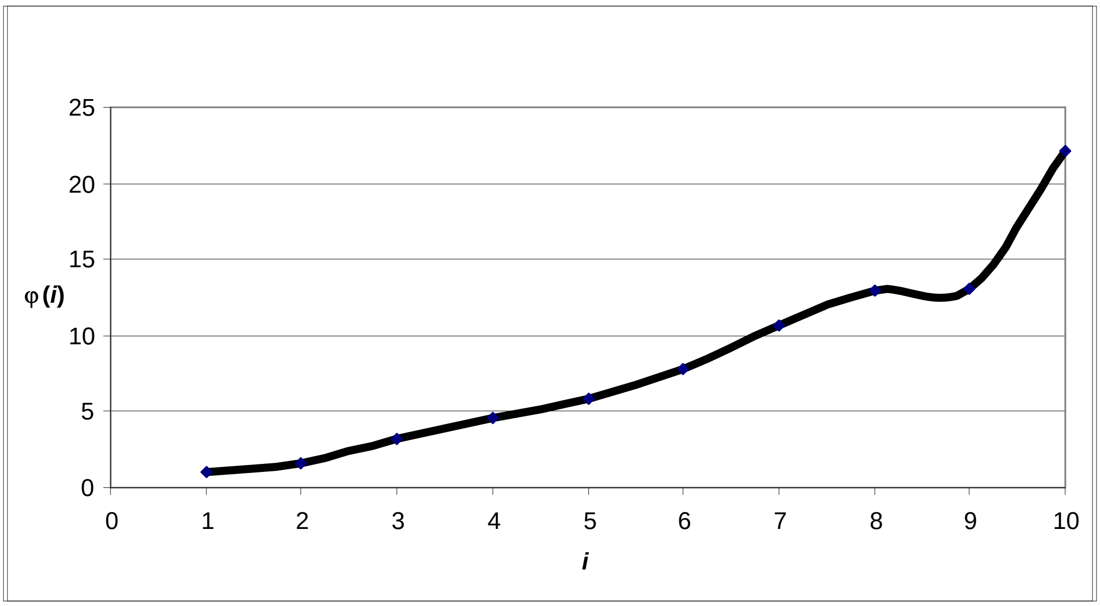
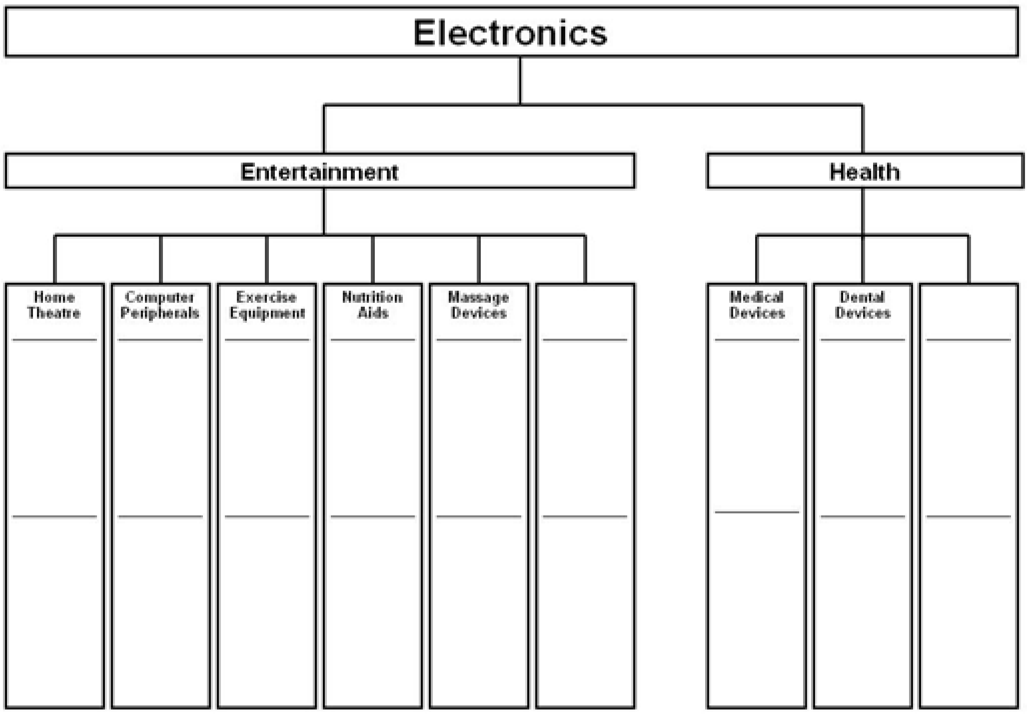


FIGURE 2

STUDY 1 CATEGORY STRUCTURES BY CONDITION

Entertainment Subcategory Numerosity Condition



Health Subcategory Numerosity Condition

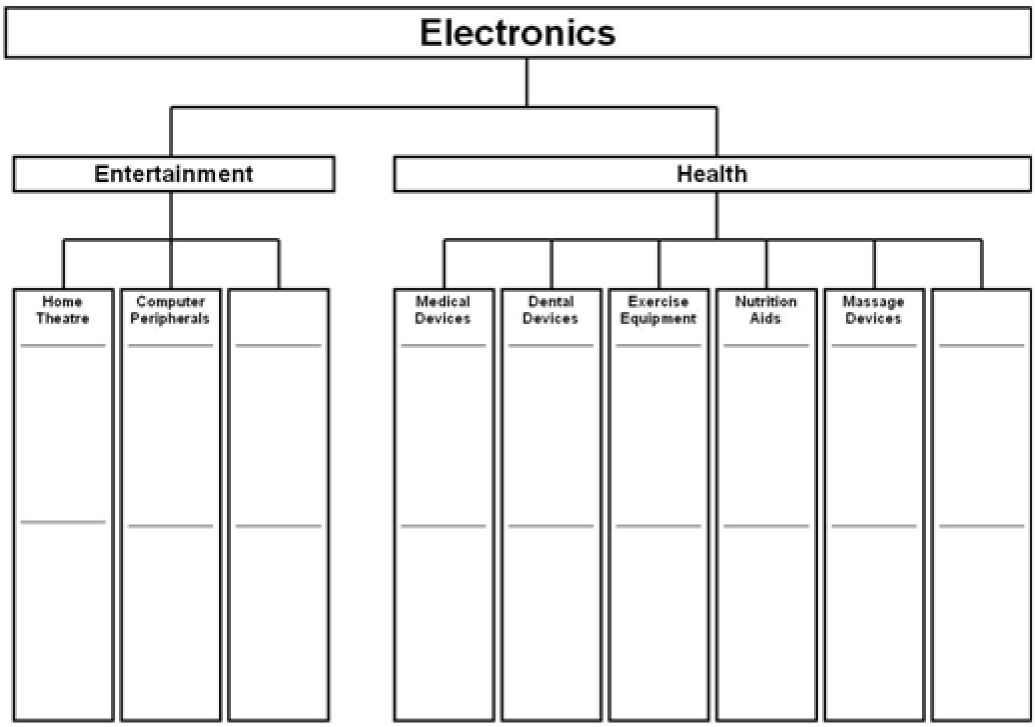


FIGURE 3
STUDY 1 CATEGORY FAMILIARIZATION PRODUCTS

Home Theatre
Kenwood Stereo System with Dolby Digital 5.1
Sony High Definition Television with Ultra Flat Screen
Computer Peripherals
Thrustmaster cockpit simulator with joystick and floor pedals
Logitech webcam with face-tracking software
Medical Devices
Welch Allyn blood pressure monitor with digital display
Minimed Insulin pump with backlit display
Dental Devices
Oral-B electric toothbrush, with floss action brush head
Phillips ultra-sonic plaque remover with adjustable intensity
Exercise Equipment
Pacific Fitness treadmill with 25% maximum incline
ProForm stationary bicycle with 30 pre-programmed courses
Nutrition Aids
Juiceman smoothie-maker with 20 speed settings
Omega electric wheatgrass juicer with twin gear technology
Massage Devices
Human Touch robotic massage recliner with heat
Sonic Skin ultrasonic face massager with anti-wrinkle cream

FIGURE 4
STUDY 1 AND 2 NEW, HYBRID PRODUCT

<u>NEW:</u> Exercise Buddy

Exercise Buddy makes working-out fun and safe:
Stores up to 1 Gb of MP3s -Jam out to your favorite music!
Built-in AM, FM, and XM radio -Tune in to your favorite programs!
Continuously monitors your heart rate, blood pressure, body temperature, and hydration levels –Keep your work-out safe!
Monitors the number of calories you have burned –Take the guesswork out of weight loss!
Connects to your computer or high definition TV via Bluetooth -Continuously monitor your vital levels!
Uploads exercise data to your computer -Track your progress!
Slim, shock resistant case –Tougher than your opponent in full-contact sports!

FIGURE 5

STUDY 2 CATEGORY FAMILIARIZATION PRODUCTS

Bicycle Computers
Filzer Bicycle Computer
VDO Wireless Bicycle Computer
Dive Computers
Suunto Vyper Dive Computer
Aladin Cobra Dive Computer
DVD Players
Peekton DVD Player Divx Multizones
H&B DVD Player
Electronic Body Scales
Tefal Electronic Body Scale
EKS White Metal Electronic Body Scale
Electronic Food Scales
Terraillon Inox Electronic Food Scale
WIK Electronic Food Scale
GPS Systems
TomTom GPS Europe
ViaMichelin GPS France
Handheld Electric Massagers
Bestron Infrared Electric Massager
OXO Design Kwaq Electric Massager
Humidifiers
White and Brown Humidifier
Tigex Humidifier
Ice Cream Makers
Bestron Ice Cream Maker
Elta Ice Cream Maker
Juicers
Riviera & Bar Juicer
Bestron Juicer
Two-Way Radios
Motorola Two-Way Radio Family Pack
Kenwood 6 Km Portable Two-Way Radios

Europe Campus
Boulevard de Constance,
77305 Fontainebleau Cedex, France
Tel: +33 (0)1 6072 40 00
Fax: +33 (0)1 60 74 00/01

Asia Campus
1 Ayer Rajah Avenue, Singapore 138676
Tel: +65 67 99 53 88
Fax: +65 67 99 53 99

www.insead.edu

INSEAD

The Business School
for the World