Concepts and Conceptual Structure

Douglas L. Medin *University of Illinois at Urbana*

ABSTRACT: Research and theory on categorization and conceptual structure have recently undergone two major shifts. The first shift is from the assumption that concepts have defining properties (the classical view) to the idea that concept representations may be based on properties that are only characteristic or typical of category examples (the probabilistic view). Both the probabilistic view and the classical view assume that categorization is driven by similarity relations. A major problem with describing category structure in terms of similarity is that the notion of similarity is too unconstrained to give an account of conceptual coherence. The second major shift is from the idea that concepts are organized by similarity to the idea that concepts are organized around theories. In this article, the evidence and rationale associated with these shifts are described, and one means of integrating similarity-based and theory-driven categorization is outlined.

What good are categories? Categorization involves treating two or more distinct entities as in some way equivalent in the service of accessing knowledge and making predictions. Take psychodiagnostic categories as an example. The need to access relevant knowledge explains why clinical psychologists do not (or could not) treat each individual as unique. Although one would expect treatment plans to be tailored to the needs of individuals, absolute uniqueness imposes the prohibitive cost of ignorance. Clinicians need some way to bring their knowledge and experience to bear on the problem under consideration, and that requires the appreciation of some similarity or relationship between the current situation and what has gone before. Although clinical psychologists may or may not use a specific categorization system, they must find points of contact between previous situations and the current context; that is, they must categorize. Diagnostic categories allow clinicians to predict the efficacy of alternative treatments and to share their experiences with other therapists. Yet another reason to categorize is to learn about etiology. People who show a common manifestation of some problem may share common precipitating conditions or causes. Ironically, the only case in which categorization would not be useful is where all individuals are treated alike; thus, categorization allows diversity.

More generally speaking, concepts and categories serve as building blocks for human thought and behavior. Roughly, a *concept* is an idea that includes all that is characteristically associated with it. A *category* is a partitioning or class to which some assertion or set of assertions might apply. It is tempting to think of categories as existing in the world and of concepts as corresponding to mental representations of them, but this analysis is misleading. It is misleading because concepts need not have real-world counterparts (e.g., unicorns) and because people may impose rather than discover structure in the world. I believe that questions about the nature of categories may be psychological questions as much as metaphysical questions. Indeed, for at least the last decade my colleagues and I have been trying to address the question of why we have the categories we have and not others. The world could be partitioned in a limitless variety of ways, yet people find only a miniscule subset of possible classifications to be meaningful. Part of the answer to the categorization question likely does depend on the nature of the world, but part also surely depends on the nature of the organism and its goals. Dolphins have no use for psychodiagnostic categories.

Given the fundamental character of concepts and categories, one might think that people who study concepts would have converged on a stable consensus with respect to conceptual structure. After all, Plato and Aristotle had quite a bit to say about concepts, medieval philosophers were obsessed with questions about universals and the essence of concepts, and concept representation remains as a cornerstone issue in all aspects of cognitive science. However, we have neither consensus nor stability. The relatively recent past has experienced at least one and probably two major shifts in thought about conceptual structure, and stability is the least salient attribute of the current situation. In the remainder of this article, I will briefly describe these shifts and then outline some ways of integrating the strong points of the various views.

The First Shift: Classical Versus Probabilistic Views

It is difficult to discuss concepts without bringing in the notion of similarity at some point. For example, a common idea is that our classification system tends to maximize within-category similarity relative to between-category similarity. That is, we group things into categories because they are similar. It will be suggested that alter-

 0003-066X/89/$00.75
Vol. 44, No. 12, 1469-1481

native views of conceptual structure are associated with distinct (though sometimes implicit) theories of the nature of similarity.

The Classical View

The idea that all instances or examples of a category have some fundamental characteristics in common that determine their membership is very compelling. The classical view of concepts is organized around this notion. The classical view assumes that mental representations of categories consist of summary lists of features or properties that individually are necessary for category membership and collectively are sufficient to determine category membership. The category *triangle* meets these criteria. All triangles are closed geometric forms with three sides and interior angles that sum to 180 degrees. To see if something is a triangle one has only to check for these three properties, and if any one is missing one does not have a triangle.

What about other concepts? The classical view suggests that all categories have defining features. A particular person may not know what these defining features are but an expert certainly should. In our 1981 book, *Categories and Concepts,* Ed Smith and I reviewed the status of the classical view as a theory of conceptual structure. We concluded that the classical view was in grave trouble for a variety of reasons. Many of the arguments and counterarguments are quite detailed, but the most serious problems can be easily summarized:

1. Failure to specify defining features. One glaring problem is that even experts cannot come up with defining features for most lexical concepts (i.e., those reflected in our language). People may believe that concepts have necessary or sufficient features (McNamara & Sternberg, 1983), but the features given as candidates do not hold up to closer scrutiny. For example, a person may list "made of wood" as a necessary property for violins, but not all violins are made of wood. Linguists, philosophers, biologists, and clinical psychologists alike have been unable to supply a core set of features that all examples of a concept (in their area of expertise) necessarily must share.

2. Goodness of example effects. According to the classical view, all examples of a concept are equally good because they all possess the requisite defining features. Experience and (by now) a considerable body of research undermines this claim. For example, people judge a robin to be a better example of bird than an ostrich is and can answer category membership questions more quickly for good examples than for poor examples (Smith, Shoben, & Rips, 1974). Typicality effects are nearly ubiquitous (for reviews, see Medin & Smith, 1984; Mervis & Rosch, 1981; Oden, 1987); they hold for the artistic style (Hartley & Homa, 1981), chess (Goldin, 1978), emotion terms (Fehr, 1988; Fehr & Russell, 1984), medical diagnosis (Arkes & Harkness, 1980), and person perception (e.g., Cantor & Mischel, 1977).

Typicality effects are not, in principle, fatal for the classical view. One might imagine that some signs or features help to determine the presence of other (defining) features. Some examples may have more signs or clearer signs pointing the way to the defining properties, and this might account for the difference in goodness of example judgments or response times. This distinction between identification procedures (how one identifies an instance of a concept) and a conceptual core (how the concept relates to other concepts) may prove useful if it can be shown that the core is used in some other aspect of thinking. It seems, however, that this distinction serves more to insulate the classical view from empirical findings, and Smith, Rips, and Medin (1984) argued that there are no sharp boundaries between core properties and those used for purposes of identification.

3. Unclear cases. The classical view implies a procedure for unambiguously determining category membership; that is, check for defining features. Yet there are numerous cases in which it is not clear whether an example belongs to a category. Should a rug be considered furniture? What about a clock or radio? People not only disagree with each other concerning category membership but also contradict themselves when asked about membership on separate occasions (Barsalou, 1989; Bellezza, 1984; McCloskey & Glucksberg, 1978).

These and other problems have led to disenchantment with the classical view of concepts. The scholarly consensus has shifted its allegiance to an alternative, the probabilistic view.

The Probabilistic View

The rejection of the classical view of categories has been associated with the ascendance of the probabilistic view of category structure (Wittgenstein, 1953). This view holds that categories are "fuzzy" or ill-defined and that categories are organized around a set of properties or clusters of correlated attributes (Rosch, 1975) that are only characteristic or typical of category membership. Thus, the probabilistic view rejects the notion of defining features.

The most recent edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-IIIR,* American Psychiatric Association, 1987) uses criteria based on lists of characteristic symptoms or features to describe diagnostic categories and thereby endorses the probabilistic view. For example, a diagnosis of depression can be made if a dysphoric mood and any five of a set of nine symptoms are present nearly every day for a period of at least two weeks. Thus, two people may both be categorized as depressed and share only a single one of the nine characteristic symptoms!

The probabilistic view is perfectly at home with the

The research described in this article was supported in part by National Science Foundation Grant No. BNS 84-19756 and by National Library of Medicine Grant No. LM 04375. Brian Ross, Edward Shoben, Ellen Markman, Greg Oden, and Dedre Gentner provided helpful comments on an earlier draft of the article.

Correspondence concerning this article should be addressed to Douglas L. Medin, who is now at Department of Psychology, University of Michigan, 330 Packard Road, Ann Arbor, MI 48104.

typicality effects that were so awkward for the classical view. Membership in probabilistic categories is naturally graded, rather than all or none, and the better or more typical members have more characteristic properties than the poorer ones. It is also easy to see that the probabilistic view may lead to unclear cases. Any one example may have several typical properties of a category but not so many that it clearly qualifies for category membership.

In some pioneering work aimed at clarifying the structural basis of fuzzy categories, Rosch and Mervis (1975) had subjects list properties of exemplars for a variety of concepts such as *bird, fruit,* and *tool.* They found that the listed properties for some exemplars occurred frequently in other category members, whereas others had properties that occurred less frequently. Most important, the more frequently an exemplar's properties appeared within a category, the higher was its rated typicality for that category. The correlation between number of characteristic properties possessed and typicality rating was very high and positive. For example, robins have characteristic bird properties of flying, singing, eating worms, and building nests in trees, and they are rated to be very typical birds. Penguins have none of these properties, and they are rated as very atypical birds. In short, the Rosch and Mervis work relating typicality to number of characteristic properties put the probabilistic view on fairly firm footing.

1. Mental representations of probabilistic view categories. If categories are not represented in terms of definitions, what form do our mental representations take? The term, probabilistic view, seems to imply that people organize categories via statistical reasoning. Actually, however, there is a more natural interpretation of fuzzy categories. Intuitively, probabilistic view categories are organized according to a *family resemblance* principle. A simple form of summary representation would be an example or ideal that possessed all of the characteristic features of a category. This summary representation is referred to as the *prototype,* and the prototype can be used to decide category membership. If some candidate example is similar enough to the prototype for a category, then it will be classified as a member of that category. The general notion is that, based on experience with examples of a category, people abstract out the central tendency or prototype that becomes the summary mental representation for the category.

A more radical principle of mental representation, which is also consistent with fuzzy categories, is the exemplar view (Smith & Medin, 1981). The exemplar view denies that there is a single summary representation and instead claims that categories are represented by means of examples. In this view, clients may be diagnosed as suicidal, not because they are similar to some prototype of a suicidal person, but because they remind the clinician of a previous client who was suicidal.

A considerable amount of research effort has been aimed at contrasting exemplar and prototype representations (see Allen, Brooks, Norman, & Rosenthal, 1988; Estes, 1986a, 1986b; Medin, 1986; Medin & Smith, 1984; Nosofsky, 1987, 1988a; and Oden, 1987). Genero and Cantor (1987) suggested that prototypes serve untrained diagnosticians well but that trained diagnosticians may find exemplars to be more helpful. For my present purposes, however, I will blur over this distinction to note that both prototype and exemplar theories rely on roughly the same similarity principle. That is, category membership is determined by whether some candidate is sufficiently similar either to the prototype or to a set of encoded examples, where similarity is based on matches and mismatches of independent, equally abstract, features.

2. Probabilistic view and similarity. To give meaning to the claim that categorization is based on similarity, it is important to be specific about what one means by similarity. Although the consensus is not uniform, I believe that the modal model of similarity with respect to conceptual structure can be summarized in terms of the four assumptions as follows: (a) Similarity between two things increases as a function of the number of features or properties they share and decreases as a function of mismatching or distinctive features. (b) These features can be treated as independent and additive. (c) The features determining similarity are all roughly the same level of abstractness (as a special case they may be irreducible primitives). (d) These similarity principles are sufficient to describe conceptual structure, and therefore, a concept is more or less equivalent to a list of its features. This theory of similarity is very compatible with the notion that categories are organized around prototypes. Nonetheless, I will later argue that each of these assumptions is wrong or misleading and that to understand conceptual structure theories of similarity are needed that reject each of these assumptions. Before outlining an alternative set of similarity assumptions, however, I will first describe a set of observations that motivate the second, still more recent, shift in thinking concerning conceptual structure.

Problems for Probabilistic View Theories

Problems for Prototypes

Although the general idea that concepts are organized around prototypes remains popular, at a more specific, empirical level, prototype theories have not fared very well. First of all, prototype theories treat concepts as context-independent. Roth and Shoben (1983), however, have shown that typicality judgments vary as a function of particular contexts. For example, tea is judged to be a more typical beverage than milk in the context of secretaries taking a break, but this ordering reverses for the context of truck drivers taking a break. Similarly, Shoben and I (Medin & Shoben, 1988) noted that the typicality of combined concepts cannot be predicted from the typicality of the constituents. As an illustrative example, consider the concept of *spoon.* People rate small spoons as more typical spoons than large spoons, and metal spoons as more typical spoons than wooden spoons. If the concept *spoon* is represented by a prototypic spoon, then a small metal spoon should be the most typical

spoon, followed by small wooden and large metal spoons, and large wooden spoons should be the least typical. Instead, people find large wooden spoons to be more typical spoons than either small wooden spoons or large metal spoons (see also Malt & Smith, 1983). The only way for a prototype model to handle these results is to posit multiple prototypes. But this strategy creates new problems. Obviously one cannot have a separate prototype for every adjective noun combination because there are simply too many possible combinations. One might suggest that there are distinct subtypes for concepts like *spoon,* but one would need a theory describing how and when subtypes are created. Current prototype models do not provide such a theory. A third problem for prototype theories grows out of Barsalou's work (1985, 1987) on goal-derived categories such as "things to take on a camping trip" and "foods to eat while on a diet." Barsalou has found that goal-derived categories show the same typicality effects as other categories. The basis for these effects, however, is not similarity to an average or prototype but rather similarity to an ideal. For example, for the category of things to eat while on a diet, typicality ratings are determined by how closely an example conforms to the ideal of zero calories.

Laboratory studies of categorization using artificially constructed categories also raise problems for prototypes. Normally many variables relevant to human classification are correlated and therefore confounded with one another. The general rationale for laboratory studies with artificially created categories is that one can isolate some variable or set of variables of interest and unconfound some natural correlations. Salient phenomena associated with fuzzy categories are observed with artificially constructed categories, and several of these are consistent with prototype theories. For example, one observes typicality effects in learning and on transfer tests using both correctness and reaction time as the dependent variable (e.g., Rosch & Mervis, 1975). A striking phenomenon, readily obtained, is that the prototype for a category may be classified more accurately during transfer tests than are the previously seen examples that were used during original category learning (e.g., Homa & Vosburgh, 1976; Medin & Schaffer, 1978; Peterson, Meagher, Chait, & Gillie, 1973).

Typicality effects and excellent classification of prototypes are consistent with the idea that people are learning these ill-defined categories by forming prototypes. More detailed analyses, however, are more problematic. Prototype theory implies that the only information abstracted from categories is the central tendency. A prototype representation discards information concerning category size, the variability of the examples, and information concerning correlations of attributes. The evidence suggests that people are sensitive to all three of these types of information (Estes, 1986b; Flannagan, Fried, & Holyoak, 1986; Fried & Holyoak, 1984; Medin, Altom, Edelson, & Freko, 1982; Medin & Schaffer, 1978). An example involving correlated attributes pinpoints part of the problem. Most people have the intuition that small birds are much more likely to sing than large birds. This intuition cannot be obtained from a single summary prototype for birds. The fact that one can generate large numbers of such correlations is a problem for the idea that people reason using prototypes. More generally, prototype representations seem to discard too much information that can be shown to be relevant to human categorizations.

Yet another problem for prototypes is that they make the wrong predictions about which category structures should be easy or difficult to learn. One way to conceptualize the process of classifying examples on the basis of similarity to prototypes is that it involves a summing of evidence against a criterion. For example, if an instance shows a criterial sum of features (appropriately weighted), then it will be classified as a bird, and the more typical a member is of the category, the more quickly the criterion will be exceeded. The key aspect of this prediction is that there must exist some additive combination of properties and their weights that can be used to correctly assign instances as members or nonmembers. The technical term for this constraint is that categories must be linearly separable (Sebestyn, 1962). For a prototype process to work in the sense of accepting all members and rejecting all nonmembers, the categories must be linearly separable.

If linear separability acts as a constraint on human categorization, then with other factors equal, people should find it easier to learn categories that are linearly separable than categories that are not linearly separable. To make a long story short, however, studies employing a variety of stimulus materials, category sizes, subject populations, and instructions have failed to find any evidence that linear separability acts as a constraint on human classification learning (Kemler-Nelson, 1984; Medin & Schwanenflugel, 1981; see also Shepard, Hovland, & Jenkins, 1961).

The cumulative effect of these various chunks of evidence has been to raise serious questions concerning the viability of prototype theories. Prototype theories imply constraints that are not observed in human categorization, predict insensitivity to information that people readily use, and fail to reflect the context sensitivity that is evident in human categorization. Rather than getting at the character of human conceptual representation, prototypes appear to be more of a caricature of it. Exemplar models handle some of these phenomena, but they fail to address some of the most fundamental questions concerning conceptual structure.

Exemplar-Based Theories

The problems just described hold not only for prototype theories in particular but also for any similarity-based categorization model that assumes that the constituent features are independent and additive. To give but one example, one could have an exemplar model of categorization that assumes that, during learning, people store examples but that new examples are classified by "computing" prototypes and determining the similarity of the novel example to the newly constructed prototypes. In

short, the central tendency would be abstracted (and other information discarded) at the time of retrieval rather than at the time of storage or initial encoding. Such a model would inherit all the shortcomings of standard prototype theories.

Some exemplar storage theories do not endorse the notion of feature independence (Hintzman, 1986; Medin & Schaffer, 1978), or they assume that classification is based on retrieving only a subset of the stored examples (presumably the most similar ones or, as a special case, the most similar one). The idea that retrieval is limited, similarity-based, and context-sensitive is in accord with much of the memory literature (e.g., Tulving, 1983). In addition, these exemplar models predict sensitivity to category size, instance variability, context, and correlated attributes. It is my impression that in head-to-head competition, exemplar models have been substantially more successful than prototype models (Barsalou & Medin, 1986; Estes, 1986b; Medin & Ross, 1989; Nosofsky, 1988a, 1988b; but see Homa, 1984, for a different opinion).

Why should exemplar models fare better than prototype models? One of the main functions of classification is that it allows one to make inferences and predictions on the basis of partial information (see Anderson, 1988). Here I am using classification loosely to refer to any means by which prior (relevant) knowledge is brought to bear, ranging from a formal classification scheme to an idiosyncratic reminding of a previous case (which, of course, is in the spirit of exemplar models; see also Kolodner, 1984). In psychotherapy, clinicians are constantly making predictions about the likelihood of future behaviors or the efficacy of a particular treatment based on classification. Relative to prototype models, exemplar models tend to be conservative about discarding information that facilitates predictions. For instance, sensitivity to correlations of properties within a category enables finer predictions: From noting that a bird is large, one can predict that it cannot sing. It may seem that exemplar models do not discard any information at all, but they are incomplete without assumptions concerning retrieval or access. In general, however, the pairs of storage and retrieval assumptions associated with exemplar models preserve much more information than prototype models. In a general review of research on categorization and problem-solving, Brian Ross and I concluded that abstraction is both conservative and tied to the details of specific examples in a manner more in the spirit of exemplar models than prototype models (Medin & Ross, 1989).

Unfortunately, context-sensitive, conservative categorization is not enough. The debate between prototype and exemplar models has taken place on a platform constructed in terms of similarity-based categorization. The second shift is that this platform has started to crumble, and the viability of probabilistic view theories of categorization is being seriously questioned. There are two central problems. One is that probabilistic view theories do not say anything about why we have the categories we have. This problem is most glaringly obvious for exemplar models that appear to allow any set of examples to form a category. The second central problem is with the notion of similarity. Do things belong in the same category because they are similar, or do they seem similar because they are in the same category?

Does Similarity Explain Categorization?

1. Flexibility. Similarity is a very intuitive notion. Unfortunately, it is even more elusive than it is intuitive. One problem with using similarity to define categories is that similarity is too flexible. Consider, for example, Tversky's (1977) influential contrast model, which defines similarity as a function of common and distinctive features weighted for salience or importance. According to this model, similarity relationships will depend heavily on the particular weights given to individual properties or features. For example, a *zebra* and a *barber pole* would be more similar than a *zebra* and a *horse* if the feature "striped" had sufficient weight. This would not necessarily be a problem if the weights were stable. However, Tversky and others have convincingly shown that the relative weighting of a feature (as well as the relative importance of matching and mismatching features) varies with the stimulus context, experimental task (Gati & Tversky, 1984; Tversky, 1977), and probably even the concept under consideration (Ortony, Vondruska, Foss, & Jones, 1985). For example, common properties shared by a pair of entities may become salient only in the context of some third entity that does not share these properties.

Once one concedes that similarity is dynamic and depends on some (not well-understood) processing principles, earlier work on the structural underpinnings of fuzzy categories can be seen in a somewhat different light. Recall that the Rosch and Mervis (1975) studies asked subjects to list attributes or properties of examples and categories. It would be a mistake to assume that people had the ability to read and report their mental representations of concepts in a veridical manner. Indeed Keil (1979, 1981) pointed out that examples like *robin* and *squirrel* shared many important properties that almost never show up in attribute listings (e.g., has a heart, breathes, sleeps, is an organism, is an object with boundaries, is a physical object, is a thing, can be thought about, and so on). In fact, Keil argued that knowledge about just these sorts of predicates, referred to as ontological knowledge (Sommers, 1971), serves to organize children's conceptual and semantic development. For present purposes, the point is that attribute listings provide a biased sample of people's conceptual knowledge. To take things a step further, one could argue that without constraints on what is to count as a feature, any two things may be arbitrarily similar or dissimilar. Thus, as Murphy and I (Murphy & Medin, 1985) suggested, the number of properties that plums and lawn mowers have in common could be infinite: Both weigh less than 1000 Kg, both are found on earth, both are found in our solar system, both cannot hear well, both have an odor, both are not worn by elephants, both are used by people, both can be dropped,

and so on (see also Goodman, 1972; Watanabe, 1969). Now consider again the status of attribute listings. They represent a biased subset of stored or readily inferred knowledge. The correlation of attribute listings with typicality judgments is a product of such knowledge and a variety of processes that operate on it. Without a theory of that knowledge and those processes, it simply is not clear what these correlations indicate about mental representations.

The general point is that attempts to describe category structure in terms of similarity will prove useful only to the extent that one specifies which principles determine what is to count as a relevant property and which principles determine the importance of particular properties. It is important to realize that the explanatory work is being done by the principles which specify these constraints rather than the general notion of similarity. In that sense similarity is more like a dependent variable than an independent variable.

2. Attribute matching and categorization. The modal model of similarity summarized in Table 1 invites one to view categorization as attribute matching. Although that may be part of the story, there are several ways in which the focus on attribute matching may be misleading. First of all, as Armstrong, Gleitman, and Gleitman (1983) emphasized, most concepts are not a simple sum of independent features. The features that are characteristically associated with the concept *bird* are just a pile of bird features unless they are held together in a "bird structure." Structure requires both attributes and *relations* binding the attributes together. Typical bird features (laying eggs, flying, having wings and feathers, building nests in trees, and singing) have both an internal structure and an external structure based on interproperty relationships. Building nests is linked to laying eggs, and building nests in trees poses logistical problems whose solution involves other properties such as having wings, flying, and singing. Thus, it makes sense to ask why birds have certain features (e.g., wings and feathers). Although people may not have thought about various interproperty relationships, they can readily reason with them. Thus, one can answer the question of why birds have wings and feathers (i.e., to fly).

In a number of contexts, categorization may be more like problem solving than attribute matching. Inferences and causal attributions may drive the categorization process. Borrowing again from work by Murphy and me (1985), "jumping into a swimming pool with one's clothes on" in all probability is not associated directly with the concept *intoxicated.* However, observing this behavior might lead one to classify the person as drunk. In general, real world knowledge is used to reason about or explain properties, not simply to match them. For example, a teenage boy might show many of the behaviors associated with an eating disorder, but the further knowledge that the teenager is on the wrestling team and trying to make a lower weight class may undermine any diagnosis of a disorder.

3. Summary. It does not appear that similarity, at least in the form it takes in current theories, is going to be at all adequate to explain categorization. Similarity may be a byproduct of conceptual coherence rather than a cause. To use a rough analogy, winning basketball teams have in common scoring more points than their opponents, but one must turn to more basic principles to explain why they score more points. One candidate for a set of deeper principles is the idea that concepts are organized around theories, and theories provide conceptual coherence. In the next section, I will briefly summarize some of the current work on the role of knowledge structures and theories in categorization and then turn to a form of rapprochement between similarity and knowledge-based categorization principle.

The Second Shift: Concepts as Organized by Theories

Knowledge-Based Categorization

It is perhaps only a modest exaggeration to say that similarity gets at the shadow rather than the substance of concepts. Something is needed to give concepts life, coherence, and meaning. Although many philosophers of science have argued that observations are necessarily theory-labeled, only recently have researchers begun to stress that the organization of concepts is knowledge-based and driven by theories about the world (e.g., Carey, 1985; S. Gelman, 1988; S. Gelman & Markman, 1986a, 1986b; Keil, 1986, 1987; Keil & Kelly, 1987; Lakoff, 1987; Markman, 1987; Massey & R. Gelman, 1988; Murphy & Medin, 1985; Oden, 1987; Rips, 1989; Schank, Collins, & Hunter, 1986; and others).

The primary differences between the similarity-based and theory-based approaches to categorization are summarized in Table 1, taken from Murphy and Medin (1985). Murphy and Medin suggested that the relation between a concept and an example is analogous to the relation between theory and data. That is, classification is not simply based on a direct matching of properties of the concept with those in the example, but rather requires that the example have the right "explanatory relationship" to the theory organizing the concept. In the case of a person diving into a swimming pool with his or her clothes on, one might try to reason back to either causes or predisposing conditions. One might believe that having too much to drink impairs judgment and that going into the pool shows poor judgment. Of course, the presence of other information, such as the fact that another person who cannot swim has fallen into the pool, would radically change the inferences drawn and, as a consequence, the categorization judgment.

One of the more promising aspects of the theory-based approach is that it begins to address the question of why we have the categories we have or why categories are sensible. In fact, coherence may be achieved in the absence of any obvious source of similarity among examples. Consider the category comprised of children, money, photo albums, and pets. Out of context the category seems odd. If one's knowledge base is enriched to

Table 1
Comparison of Two Approaches to Concepts

Aspect of conceptual theory	Similarity-based approach	Theory-based approach
Concept representation	Similarity structure, attribute lists, correlated attributes	Correlated attributes plus underlying principles that determine which correlations are noticed
Category definition	Various similarity metrics, summation of attributes	An explanatory principle common to category members
Units of analysis	Attributes	Attributes plus explicitly represented relations of attributes and concepts
Categorization basis	Attribute matching	Matching plus inferential processes supplied by underlying principles
Weighting of attributes	Cue validity, salience	Determined in part by importance in the underlying principles
Interconceptual structure	Hierarchy based on shared attributes	Network formed by causal and explanatory links, as well as sharing of properties picked out as relevant
Conceptual development	Feature accretion	Changing organization and explanations of concepts as a result of world knowledge

include the fact that the category represents "things to take out of one's house in case of a fire," the category becomes sensible (Barsalou, 1983). In addition, one could readily make judgments about whether new examples (e.g., personal papers) belonged to the category, judgments that would not be similarity based.

Similarity effects can be overridden by theory-related strategies even in the judgments of young children. That fact was very nicely demonstrated by Gelman and Markman (1986a) in their studies of induction. Specifically, they pitted category membership against perceptual similarity in an inductive inference task. Young children were taught that different novel properties were true of two examples and then were asked which property was also true of a new example that was similar to one alternative but belonged to a different category, and one that was perceptually different from the other examples but belonged to the same category. For example, children might be taught that a (pictured) flamingo feeds its baby mashed-up food and that a (pictured) bat feeds its baby milk, and then they might be asked how a (pictured) owl feeds its baby. The owl was more perceptually similar to the bat than to the flamingo, but even four-year-olds made inferences on the basis of category membership rather than similarity.

Related work by Susan Carey and Frank Keil shows that children's biological theories guide their conceptual development. For example, Keil has used the ingenious technique of describing transformations or changes such as painting a horse to look like a zebra to examine the extent to which category membership judgments are controlled by superficial perceptual properties. Biological theories determine membership judgments quite early on (Keil, 1987; Keil & Kelly, 1987). Rips (1989) has used the same technique to show that similarity is neither necessary nor sufficient to determine category membership. It even appears to be the case that theories can affect judgments of similarity. For example, Medin and Shoben (1988) found that the terms *white hair* and *grey hair* were judged to be more similar than *grey hair* and *black hair,* but that the terms *white clouds* and *grey clouds* were judged as less similar than *grey clouds* and *black clouds.* Our interpretation is that white and grey hair are linked by a theory (of aging) in a way that white and grey clouds are not.

The above observations are challenging for defenders of the idea that similarity drives conceptual organization. In fact, one might wonder if the notion of similarity is so loose and unconstrained that we might be better off without it. Goodman (1972) epitomized this attitude by calling similarity "a pretender, an imposter, a quack" (p. 437). After reviewing some reasons to continue to take similarity seriously, I outline one possible route for integrating similarity-based and theory-based categorization.

The Need for Similarity

So far I have suggested that similarity relations do not provide conceptual coherence but that theories do. Because a major problem with similarity is that it is so unconstrained, one might ask what constrains theories. If we cannot identify constraints on theories, that is, say something about why we have the theories we have and not others, then we have not solved the problem of coherence: It simply has been shifted to another level. Although I believe we can specify some general properties of theories and develop a psychology of explanation (e.g., Abelson & Lalljee, 1988; Einhorn & Hogarth, 1986; Hilton & Slugoski, 1986; Leddo, Abelson, & Gross, 1984), I equally believe that a constrained form of similarity will play an important role in our understanding of human

concepts. This role is not to provide structure so much as it is to guide learners toward structure.

The impact of more direct perceptual similarity on the development of causal explanations is evident in the structure of people's naive theories. Frazer's (1959) cross-cultural analysis of belief systems pointed to the ubiquity of two principles, homeopathy and contagion. The principle of homeopathy is that causes and effects tend to be similar. One manifestation of this principle is homeopathic medicine, in which the cure (and the cause) are seen to resemble the symptoms. In the Azande culture, for example, the cure for ringworm is to apply fowl's excrement because the excrement looks like the ringworm. Schweder (1977) adduced strong support for the claim that resemblance is a fundamental conceptual tool of everyday thinking in all cultures, not just so-called primitive cultures.

Contagion is the principle that a cause must have some form of contact to transmit its effect. In general, the more contiguous (temporally and spatially similar) events are in time and space, the more likely they are to be perceived as causally related (e.g., Dickinson, Shanks, & Evenden, 1984; Michotte, 1963). People also tend to assume that causes and effects should be of similar magnitude. Einhorn and Hogarth (1986) pointed out that the germ theory of disease initially met with great resistance because people could not imagine how such tiny organisms could have such devastating effects.

It is important to recognize that homeopathy and contagion often point us in the right direction. Immunization can be seen as a form of homeopathic medicine that has an underlying theoretical principle to support it. My reading of these observations, however, is not that specific theoretical (causal) principles are constraining similarity but rather that similarity (homeopathy and contagion) acts as a constraint on the search for causal explanations. Even in classical conditioning studies, the similarity of the conditioned stimulus and the unconditioned stimulus can have a major influence on the rate of conditioning (Testa, 1974). Of course, similarity must itself be constrained for terms like homeopathy to have a meaning. Shortly, I will suggest some constraints on similarity as part of an effort to define a role for similarity in conceptual development.

Similarity is likely to have a significant effect on explanations in another way. Given the importance of similarity in retrieval, it is likely that explanations that are applied to a novel event are constrained by similar events and their associated explanations. For example, Read (1983) found that people may rely on single, similar instances in making causal attributions about behaviors. Furthermore, Ross (1984) and Gentner and Landers (1985) have found that superficial similarities and not just similarity with respect to deeper principles or relations play a major role in determining the remindings associated with problem solving and the use of analogy.

In brief, it seems that similarity cannot be banished from the world of theories and conceptual structures. But it seems to me that a theory of similarity is needed that is quite different in character from the one summarized in Table 1. I will suggest an alternative view of similarity and then attempt to show its value in integrating and explanation with respect to concepts.

Similarity and Theory in Conceptual Structure

A Contrasting Similarity Model

The following are key tenets of the type of similarity theory needed to link similarity with knowledge-based categorization: (a) Similarity needs to include attributes, relations, and higher-order relations. (b) Properties in general are not independent but rather are linked by a variety of interproperty relations. (c) Properties exist at multiple levels of abstraction. (d) Concepts are more than lists. Properties and relations create depth or structure. Each of the four main ideas directly conflicts with the corresponding assumption of the theory of similarity outlined earlier. In one way or another all of these assumptions are tied to structure. The general idea I am proposing is far from new. In the psychology of visual perception, the need for structural approaches to similarity has been a continuing, if not major, theme (e.g., Biederman, 1985, 1987; Palmer, 1975, 1978; Pomerantz, Sager, & Stoever, 1977). Oden and Lopes (1982) have argued that this view can inform our understanding of concepts: "Although similarity must function at some level in the induction of concepts, the induced categories are not 'held together' subjectively by the undifferentiated 'force' of similarity, but rather by structural principles" (p. 78). Nonindependence of properties and simple and higher-order relations add a dimension of depth to categorization. Depth has clear implications for many of the observations that seem so problematic for probabilistic view theories. I turn now to the question of how these modified similarity notions may link up with theory-based categorization.

Psychological Essentialism

Despite the overwhelming evidence against the classical view, there is something about it that is intuitively compelling. Recently I and my colleagues have begun to take this observation seriously, not for its metaphysical implications but as a piece of psychological data (Medin & Ortony, 1989; Medin & Wattenmaker, 1987; Wattenmaker, Nakamura, & Medin, 1988). One might call this framework "psychological essentialism." The main ideas are as follows: People act as if things (e.g., objects) have essences or underlying natures that make them the thing that they are. Furthermore, the essence constrains or generates properties that may vary in their centrality. One of the things that theories do is to embody or provide causal linkages from deeper properties to more superficial or surface properties. For example, people in our culture believe that the categories *male* and *female* are genetically determined, but to pick someone out as male or female we rely on characteristics such as hair length, height, facial hair, and clothing that represent a mixture of secondary sexual characteristics and cultural conventions. Although

these characteristics are more unreliable than genetic evidence, they are far from arbitrary. Not only do they have some validity in a statistical sense, but also they are tied to our biological and cultural conceptions of *male* and *female*.

It is important to note that psychological essentialism refers not to how the world is but rather to how people approach the world. Wastebaskets probably have no true essence, although we may act as if they do. Both social and psychodiagnostic categories are at least partially culture specific and may have weak if any metaphysical underpinnings (see also Morey & McNamara, 1987).

If psychological essentialism is bad metaphysics, why should people act as if things had essences? The reason is that it may prove to be good epistomology. One could say that people adopt an *essentialist heuristic*, namely, the hypothesis that things that look alike tend to share deeper properties (similarities). Our perceptual and conceptual systems appear to have evolved such that the essentialist heuristic is very often correct (Medin & Wattenmaker, 1987; Shepard, 1984). This is true even for human artifacts such as cars, computers, and camping stoves because structure and function tend to be correlated. Surface characteristics that are perceptually obvious or are readily produced on feature listing tasks may not so much constitute the core of a concept as point toward it. This observation suggests that classifying on the basis of similarity will be relatively effective much of the time, but that similarity will yield to knowledge of deeper principles. Thus, in the work of Gelman and Markman (1986a) discussed earlier, category membership was more important than perceptual similarity in determining inductive inferences.

Related Evidence

The contrasting similarity principles presented earlier coupled with psychological essentialism provide a framework for integrating knowledge-based and similarity-based categorization. Although it is far short of a formal theory, the framework provides a useful perspective on many of the issues under discussion in this article.

1. Nonindependence of features. Earlier I mentioned that classifying on the basis of similarity to a prototype was functionally equivalent to adding up the evidence favoring a classification and applying some criterion (at least X out of Y features). Recall also that the data ran strongly against this idea. From the perspective currently under consideration, however, there ought to be two ways to produce data consistent with prototype theory. One would be to provide a theory that suggests the prototype as an ideal or that makes summing of evidence more natural. For example, suppose that the characteristic properties for one category were as follows: It is made of metal, has a regular surface, is of medium size, and is easy to grasp. For a contrasting category the characteristic properties were: It is made of rubber, has an irregular surface, is of small size, and is hard to grasp. The categories may not seem sensible or coherent but suppose one adds the information that the objects in one category could serve as substitutes for a hammer. Given this new information, it becomes easy to add up the properties of examples in terms of their utility in supporting hammering. In a series of studies using the above descriptions and related examples, Wattenmaker, Dewey, Murphy, and I (1986) found data consistent with prototype theory when the additional information was supplied, and data inconsistent with prototype theory when only characteristic properties were supplied. Specifically, they found that linearly separable categories were easier to learn than nonlinearly separable categories only when an organizing theme was provided (see also Nakamura, 1985).

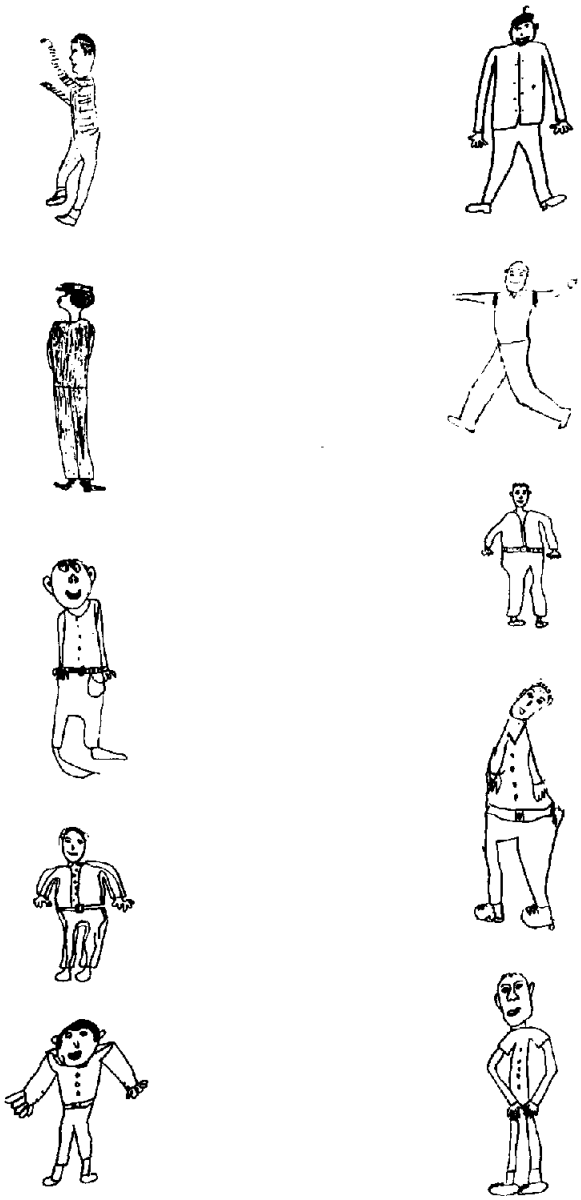
One might think that prototypes become important whenever the categories are meaningful. That is not the case. When themes are provided that are not compatible with a summing of evidence, the data are inconsistent with prototype theories. For instance, suppose that the examples consisted of descriptions of animals and that the organizing theme was that one category consisted of prey and the other of predators. It is a good adaptation for prey to be armored and to live in trees, but an animal that is both armored and lives in trees may not be better adapted than an animal with either characteristic alone. Being armored and living in trees may be somewhat incompatible. Other studies by Wattenmaker et al. using directly analogous materials failed to find any evidence that linear separability (and, presumably, summing of evidence) was important or natural. Only some kinds of interproperty relations are compatible with a summing of evidence, and evidence favoring prototypes may be confined to these cases.

The above studies show that the ease or naturalness of classification tasks cannot be predicted in terms of abstract category structures based on distribution of features, but rather requires an understanding of the knowledge brought to bear on them, for this knowledge determines inter-property relationships. So far only a few types of interproperty relationships have been explored in categorization, and much is to be gained from the careful study of further types of relations (e.g., see Barr & Caplan, 1987; Chaffin & Hermann, 1987; Rips & Conrad, 1989; Winston, Chaffin, & Herman, 1987).

2. Levels of features. Although experimenters can often contrive to have the features or properties comprising stimulus materials at roughly the same level of abstractness, in more typical circumstances levels may vary substantially. This fact has critical implications for descriptions of category structure (see Barsalou & Billman, 1988). This point may be best represented by an example from some ongoing research I am conducting with Glenn Nakamura and Ed Wisniewski. Our stimulus materials consist of children's drawings of people, a sample of which is shown in Figure 1. There are two sets of five drawings, one on the left and one on the right. The task of the participants in this experiment is to come up with a rule that could be used to correctly classify both these drawings and new examples that might be presented later.

One of our primary aims in this study was to ex-

Figure 1
Children's Drawings of People Used in the Rule Induction Studies by Nakamura, Wisniewski, and Medin



amine the effects of different types of knowledge structures on rule induction. Consequently, some participants were told that one set was done by farm children and the other by city children; some were told that one set was done by creative children and the other by noncreative children; and still others were told that one set was done by emotionally disturbed children and the other by mentally healthy children. The exact assignment of drawings was counterbalanced with respect to the categories such that half the time the drawings on the left of Figure 1 were labeled as done by farm children and half the time the drawings on the right were labeled as having been done by farm children.

Although we were obviously expecting differences in the various conditions, in some respects the most striking result is one that held across conditions. Almost without exception the rules that people gave had properties at two or three different levels of abstractness. For example, one person who was told the drawings on the left were done by city children gave the following rule: "The city drawings use more profiles, and are more elaborate. The clothes are more detailed, showing both pockets and buttons, and the hair is drawn in. The drawings put less emphasis on proportion and the legs and torso are off." Another person who was told the same drawings were done by farm children wrote: "The children draw what they see in their normal life. The people have overalls on and some drawings show body muscles as a result of labor. The drawings are also more detailed. One can see more facial details and one drawing has colored the clothes and another one shows the body under the clothes." As one can see, the rules typically consist of a general assertion or assertions coupled with either an operational definition or examples to illustrate and clarify the assertion. In some cases these definitions or examples extend across several levels of abstractness.

One might think that our participants used different levels of description because there was nothing else for them to do. That is, there may have been no low-level perceptual features that would separate the groups. In a followup study we presented examples one at a time and asked people to give their rule after each example. If people are being forced to use multiple levels of description because simple rules will not work, then we should observe a systematic increase in the use of multiple levels across examples. In fact, however, we observed multiple levels of description as the predominant strategy from the first example on. We believe that multiple levels arise when people try to find a link between abstract explanatory principles or ideas (drawings reflect one's experience) and specific details of drawings.

There are several important consequences of multilevel descriptions. First of all, the relation across levels is not necessarily a subset, superset, or a part–whole relation. Most of the time one would say that the lower level property "supports" the higher level property; for example, "jumping into a swimming pool with one's clothes on" supports poor judgment. This underlines the point that categorization often involves more than a simple matching of properties. A related point is that features are ambiguous in the sense that they may support more than one higher level property. When the drawings on the right were associated with the label *mentally healthy,* a common description was "all the faces are smiling." When the label for the same drawing was *noncreative,* a common description was "the faces show little variability in expression." Finally, it should be obvious that whether a category description is disjunctive (e.g., pig's nose or cow's mouth or catlike ears) or conjunctive or defining

(e.g., all have animal parts) depends on the level with respect to which the rule is evaluated.

3. Centrality. If properties are at different levels of abstraction and linked by a variety of relations, then one might imagine that some properties are more central than others because of the role they play in conceptual structure. An indication that properties differ in their centrality comes from a provocative study by Asch and Zukier (1984). They presented people with trait terms that appeared to be contradictory (e.g., kind and vindictive) and asked participants if these descriptions could be resolved (e.g., how could a person be both kind and vindictive?). Participants had no difficulty integrating the pairs of terms, and Asch and Zukier identified seven major resolution strategies. For present purposes, what is notable is that many of the resolution strategies involve making one trait term more central than the other one. For example, one way of integrating *kind* and *vindictive* was to say that the person was fundamentally evil and was kind only in the service of vindictive ends.

In related work, Shoben and I (Medin & Shoben, 1988) showed that centrality of a property depends on the concept of which it is a part. We asked participants to judge the typicality of adjective noun pairs when the adjective was a property that other participants judged was not true of the noun representing the concept. For example, our participants judged that all bananas and all boomerangs are curved. Based on this observation, other participants were asked to judge the typicality of a straight banana as a banana or a straight boomerang as a boomerang. Other instances of the 20 pairs used include *soft knife* versus *soft diamond* and *polka dot fire hydrant* versus *polka dot yield sign.* For 19 of the 20 pairs, participants rated one item of a pair as more typical than the other. Straight banana, soft knife, and polka dot fire hydrant were rated as more typical than straight boomerang, soft diamond, and polka dot yield sign. In the case of boomerangs (and probably yield signs), centrality may be driven by structure-function correlations. Soft diamonds are probably rated as very atypical because hardness is linked to many other properties and finding out that diamonds were soft would call a great deal of other knowledge into question.

Most recently, Woo Kyoung Ahn, Joshua Rubenstein, and I have been interviewing clinical psychologists and psychiatrists concerning their understanding of psychodiagnostic categories. Although our project is not far enough along to report any detailed results, it is clear that the *DSM-IIIR* guidebook (American Psychiatric Association, 1987) provides only a skeletal outline that is brought to life by theories and causal scenarios underlying and intertwined with the symptoms that comprise the diagnostic criteria. Symptoms differ in the level of abstractness and the types and number of intersymptom relations in which they participate, and as a consequence, they differ in their centrality.

Conclusions

The shift to a focus on knowledge-based categorization does not mean that the notion of similarity must be left behind. But we do need an updated approach to, and interpretation of, similarity. The mounting evidence on the role of theories and explanations in organizing categories is much more compatible with features at varying levels linked by a variety of interproperty relations than it is with independent features at a single level. In addition, similarity may not so much constitute structure as point toward it. There is a dimension of depth to categorization. The conjectures about psychological essentialism may be one way of reconciling classification in terms of perceptual similarity or surface properties with the deeper substance of knowledge-rich, theory-based categorization.

REFERENCES

Abelson, R. P., & Lalljee, M. G. (1988). Knowledge-structures and causal explanation. In D. J. Hilton (Ed.), *Contemporary science and natural explanation: Commonsense conceptions of causality* (pp. 175–202). Brighton, England: Harvester Press.

Allen, S. W., Brooks, L. R., Norman, G. R., & Rosenthal, D. (1988, November). *Effect of prior examples on rule-based diagnostic performance.* Paper presented at the meeting of the Psychonomic Society, Chicago.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (rev. ed.). Washington, DC: Author.

Anderson, J. R. (1988). The place of cognitive architectures in a rational analyses. In *The Tenth Annual Conference of the Cognitive Science Society* (pp. 1–10). Montreal, Canada: University of Montreal.

Arkes, H. R., & Harkness, A. R. (1980). Effect of making a diagnosis on subsequent recognition of symptoms. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 568–575.

Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). What some concepts might not be. *Cognition, 13,* 263–308.

Asch, S. E., & Zukier, H. (1984). Thinking about persons. *Journal of Personality and Social Psychology, 46,* 1230–1240.

Barr, R. A., & Caplan, L. J. (1987). Category representations and their implications for category structure. *Memory and Cognition, 15,* 397–418.

Barsalou, L. W. (1983). Ad hoc categories. *Memory and Cognition, 11,* 211–227.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory and Cognition, 11,* 629–654.

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development: The ecological and intellectual factors in categorization* (pp. 101–140). Cambridge, England: Cambridge University Press.

Barsalou, L. W. (1989). Intra-concept similarity and its implications for inter-concept similarity. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 76–121). Cambridge, England: Cambridge University Press.

Barsalou, L. W., & Billman, D. (1988, April). *Systematicity and semantic ambiguity.* Paper presented at a workshop on semantic ambiguity, Adelphi University.

Barsalou, L. W., & Medin, D. L. (1986). Concepts: Fixed definitions or dynamic context-dependent representations? *Cahiers de Psychologie Cognitive, 6,* 187–202.

Bellezza, F. S. (1984). Reliability of retrieval from semantic memory: Noun meanings. *Bulletin of the Psychonomic Society, 22,* 377–380.

Biederman, I. (1985). Human image understanding: Recent research and a theory. *Computer Vision, Graphics, and Image Processing, 32,* 29–83.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94,* 115–147.

Cantor, N., & Mischel, W. (1977). Traits as prototypes: Effects on recognition memory. *Journal of Personality and Social Psychology, 35,* 38–48.

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: Massachusetts Institute of Technology Press.

Chaffin, R., & Herrmann, D. J. (1987). Relation element theory: A new account of the representation and processing of semantic relations. In D. Gorfein & R. Hoffman (Eds.), *Learning and memory: The Ebbinghaus centennial conference* (pp. 221–245). Hillsdale, NJ: Erlbaum.

Dickinson, A., Shanks, D., & Evenden, J. (1984). Judgment of act-outcomes contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology, 36A*(1), 29–50.

Einhorn, J. H., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin, 99,* 3–19.

Estes, W. K. (1986a). Memory storage and retrieval processes in category learning. *Journal of Experimental Psychology: General, 115,* 155–175.

Estes, W. K. (1986b). Array models for category learning. *Cognitive Psychology, 18,* 500–549.

Fehr, B. (1988). Prototype analysis of the concepts of love and commitment. *Journal of Personality and Social Psychology, 55,* 557–579.

Fehr, B., & Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General, 113,* 464–486.

Flannagan, M. J., Fried, L. S., & Holyoak, K. J. (1986). Distributional expectations and the induction of category structure. *Journal of Experimental Psychology: Learning, Memory and Cognition, 12,* 241–256.

Frazer, J. G. (1959). *The new golden bough.* New York: Criterion Books.

Fried, L. S., & Holyoak, K. J. (1984). Induction of category distribution: A framework for classification learning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 10,* 234–257.

Gati, I., & Tversky, A. (1984). Weighting common and distinctive features in perceptual and conceptual judgments. *Cognitive Psychology, 16,* 341–370.

Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology, 20,* 65–95.

Gelman, S. A., & Markman, E. M. (1986a). Categories and induction in young children. *Cognition, 23,* 183–209.

Gelman, S. A., & Markman, E. M. (1986b). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development, 58,* 1532–1541.

Genero, N., & Cantor, N. (1987). Exemplar prototypes and clinical diagnosis: Toward a cognitive economy. *Journal of Social and Clinical Psychology, 5,* 59–78.

Gentner, D., & Landers, R. (1985). *Analogical reminding: A good match is hard to find.* Paper presented at the International Conference of Systems, Man and Cybernetics, Tucson, AZ.

Goldin, S. E. (1978). Memory for the ordinary: Typicality effects in chess memory. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 605–616.

Goodman, N. (1972). Seven strictures on similarity. In N. Goodman (Ed.), *Problems and projects.* New York: Bobbs-Merrill.

Hartley, J., & Homa, D. (1981). Abstraction of stylistic concepts. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 33–46.

Hilton, D. J., & Slugoski, B. R. (1986). Knowledge-based causal attribution: The abnormal conditions focus model. *Psychological Review, 93,* 75–88.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93,* 411–428.

Homa, D. (1984). On the nature of categories. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 18, pp. 49–94). New York: Academic Press.

Homa, D., & Vosburgh, R. (1976). Category breadth and the abstraction of prototypical information. *JEP: Human Learning and Memory, 2,* 322–330.

Keil, F. C. (1979). *Semantic and conceptual development: An ontological perspective.* Cambridge, MA: Harvard University Press.

Keil, F. C. (1981). Constraints on knowledge and cognitive development. *Psychological Review, 88,* 197–227.

Keil, F. C. (1986). The acquisition of natural kind and artifact terms. In W. Demopoulos & A. Marras (Eds.), *Language learning and concept acquisition* (pp. 133–153). Norwood, NJ: Ablex.

Keil, F. C. (1987). Conceptual development and category structure. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 175–200). Cambridge, England: Cambridge University Press.

Keil, F. C., & Kelly, M. H. (1987). Developmental changes in category structure. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 491–510). Cambridge, England: Cambridge University Press.

Kemler-Nelson, D. G. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning and Verbal Behavior, 23,* 734–759.

Kolodner, J. L. (1984). *Retrieval and organizational structures in conceptual memory: A computer model.* Hillsdale, NJ: Erlbaum.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories tell us about the nature of thought.* Chicago: University of Chicago Press.

Leddo, J., Abelson, R. P., & Gross, P. H. (1984). Conjunctive explanation: When two explanations are better than one. *Journal of Personality and Social Psychology, 47,* 933–943.

Malt, B. C., & Smith, E. E. (1983). Correlated properties in natural categories. *Journal of Verbal Learning and Verbal Behavior, 23,* 250–269.

Markman, E. M. (1987). How children constrain the possible meanings of words. In U. Neisser (Ed.), *Concepts and conceptual development: The ecological and intellectual factors in categorization* (pp. 256–287). Cambridge, England: Cambridge University Press.

Massey, C. M., & Gelman, R. (1988). Preschoolers' ability to decide whether a photographed unfamiliar object can move itself. *Developmental Psychology, 24,* 307–317.

McCloskey, M., & Glucksberg, S. (1978). Natural categories: Well-defined or fuzzy sets? *Memory and Cognition, 6,* 462–472.

McNamara, T. P., & Sternberg, R. J. (1983). Mental models of word meaning. *Journal of Verbal Learning and Verbal Behavior, 22,* 449–474.

Medin, D. L. (1986). Commentary on "Memory storage and retrieval processes in category learning." *Journal of Experimental Psychology: General, 115*(4), 373–381.

Medin, D. L., Altom, M. W., Edelson, S. M., & Freko, D. (1982). Correlated symptoms and simulated medical classification. *Journal of Experimental Psychology: Learning, Memory and Cognition, 8,* 37–50.

Medin, D. L., & Ortony, A. (1989). *Psychological essentialism.* In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). New York: Cambridge University Press.

Medin, D. L., & Ross, B. H. (1989). The specific character of abstract thought: Categorization, problem-solving, and induction. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 5, pp. 189–223). Hillsdale, NJ: Erlbaum.

Medin, D. L., & Schaffer, M. M. (1978). A context theory of classification learning. *Psychological Review, 85,* 207–238.

Medin, D. L., & Schwanenflugel, P. J. (1981). Linear separability in classification learning. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 355–368.

Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology, 20,* 158–190.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual Review of Psychology, 35,* 113–118.

Medin, D. L., & Wattenmaker, W. D. (1987). Category cohesiveness, theories, and cognitive archeology. In U. Neisser (Ed.), *Concepts and conceptual development: The ecological and intellectual factors in categories* (pp. 25–62). Cambridge, England: Cambridge University Press.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual Review of Psychology, 32,* 89–115.

Michotte, A. (1963). *Perception of causality.* London: Methuen.

Morey, L. C., & McNamara, T. P. (1987). On definitions, diagnosis, and DSM-III. *Journal of Abnormal Psychology, 96,* 283–285.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92,* 289–316.

Nakamura, G. V. (1985). Knowledge-based classification of ill-defined categories. *Memory and Cognition, 13,* 377–384.

Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 87–108.

Nosofsky, R. M. (1988a). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 700–708.

Nosofsky, R. M. (1988b). Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 54–65.

Oden, G. C. (1987). Concept, knowledge, and thought. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual Review of Psychology, 38,* 203–227.

Oden, G. C., & Lopes, L. (1982). On the internal structure of fuzzy subjective categories. In R. R. Yager (Ed.), *Recent developments in fuzzy set and possibility theory* (pp. 75–89). Elmsford, NY: Pergamon Press.

Ortony, A., Vondruska, R. J., Foss, M. A., & Jones, L. E. (1985). Salience, similes, and the asymmetry of similarity. *Journal of Memory and Language, 24,* 569–594.

Palmer, S. E. (1975). Visual perception and world knowledge. In D. A. Norman & D. E. Rumelhart (Eds.), *Explorations in cognition* (pp. 279–307). San Francisco: W. H. Freeman.

Palmer, S. E. (1978). Structural aspects of visual similarity. *Memory and Cognition, 6,* 91–97.

Peterson, M. J., Meagher, R. B., Jr., Chait, H., & Gillie, S. (1973). The abstraction and generalization of dot patterns. *Cognitive Psychology, 4,* 378–398.

Pomerantz, J. R., Sager, L. C., & Stoever, R. G. (1977). Perception of wholes and their component parts: Some configural superiority effects. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 422–435.

Read, S. J. (1983). Once is enough: Causal reasoning from a single instance. *Journal of Personality and Social Psychology, 45,* 323–334.

Rips, L. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21–59). New York: Cambridge University Press.

Rips, L. J., & Conrad, F. G. (1989). The folk psychology of mental activities. *Psychological Review, 96,* 187–207.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104,* 192–233.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573–605.

Ross, B. H. (1984). Remindings and their effects in learning a cognitive skill. *Cognitive Psychology, 16,* 371–416.

Roth, E. M., & Shoben, E. J. (1983). The effect of context on the structure of categories. *Cognitive Psychology, 15,* 346–378.

Schank, R. C., Collins, G. C., & Hunter, L. E. (1986). Transcending induction category formation in learning. *The Behavioral and Brain Sciences, 9,* 639–686.

Schweder, R. A. (1977). Likeness and likelihood in everyday thought: Magical thinking in judgments about personality. *Current Anthropology, 18,* 4.

Sebestyn, G. S. (1962). *Decision-making processes in pattern recognition.* New York: Macmillan.

Shepard, R. H. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review, 19,* 417–447.

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs, 75,* (13, Whole No. 517).

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Smith, E. E., Rips, J. J., & Medin, D. W. (1984). A psychological approach to concepts: Comments on Rey's "Concepts and stereotypes." *Cognition, 17,* 265–274.

Smith, E. E., Shoben, E. J., & Rips, J. J. (1974). Structure and processes in semantic memory: A featural model for semantic decisions. *Psychological Review, 81,* 214–241.

Sommers, F. (1971). Structural ontology. *Philosophia, 1,* 21–42.

Testa, T. J. (1974). Causal relationships and the acquisition of avoidance responses. *Psychological Review, 81,* 491–505.

Tulving, E. (1983). *Elements of episodic memory.* New York: Oxford University Press.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327–352.

Watanabe, S. (1969). *Knowing and guessing: A formal and quantitative study.* New York: Wiley.

Wattenmaker, W. D., Dewey, G. I., Murphy, T. D., & Medin, D. L. (1986). Linear separability and concept learning: Context, relational properties, and concept naturalness. *Cognitive Psychology, 18,* 158–194.

Wattenmaker, W. D., Nakamura, G. V., & Medin, D. L. (1988). Relationships between similarity-based and explanation-based categorization. In D. Hilton (Ed.), *Contemporary science and natural explanation: Commonsense conceptions of causality* (pp. 205–241). Brighton, England: Harvester Press.

Winston, M. E., Chaffin, R., & Herrmann, D. (1987). A taxonomy of part-whole relations. *Cognitive Science, 11,* 417–444.

Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, trans.). Oxford, England: Blackwell.