Psychological Bulletin
1992, Vol. 112, No. 3, 500–526

0033-2909/92/$3.00

Recent Views of Conceptual Structure

Lloyd K. Komatsu
Carleton College

This article reviews theories of concept structure proposed since the mid-1970s, when the discovery of typicality effects led to the rejection of the view that instances of a concept share necessary and sufficient attributes. To replace that classical view, psychologists proposed the family resemblance and exemplar views (and hybrids of the 2), which argue that instances of a concept share a certain level of overall similarity, rather than necessary and sufficient attributes. These similarity-based views account for much of the typicality data but fail to provide an adequate explanation of the coherence of conceptual categories and of various context effects. Recently proposed explanation-based accounts address these issues but raise further questions about the distinction between concept-specific information and general knowledge and about the relationship between conceptual knowledge and various forms of inference.

Psychologists have traditionally equated knowing the meaning of a word with knowing (or perhaps more accurately, having) the concept labeled by a word (e.g., Ogden & Richards, 1956; but see Clark, 1983). In this approach, a concept is assumed to be the mental representation of a category or class (Gleitman, Armstrong, & Gleitman, 1983; Medin & Smith, 1984). The contents of such a mental representation (i.e., the intension of a word), in concert with certain assumptions about how those contents are processed, have been taken to explain a wide variety of phenomena, including people's knowledge of linguistic relations (e.g., synonymy, antynomy, hyponomy), how people recognize the objects, events, and so on properly labeled by the word (i.e., the extension of the word), how people understand novel combinations of the word with other words, and the inferences people are able to make about an object, event, and so on, properly labeled by the word (Johnson-Laird, Herrmann, & Chaffin, 1984; Medin & Smith, 1984; E. E. Smith & Medin, 1981).

Over the years, psychologists have described concepts in a variety of ways. Building on a convention established by E. E. Smith and Medin (1981; see also Medin & Smith, 1984), I describe psychological accounts of concepts as taking one of five views: the classical, the family resemblance (or probablistic), the exemplar, the schema, and the explanation based. Following Murphy and Medin (1985), I refer to the classical, family resemblance, and exemplar views as being similarity based, to contrast them with the more recent explanation-based view.

To take a concrete example, what information, very generally, is represented by the concept chair, so that people are able to reason about chairs, recognize instances of chairs, and understand complex concepts such as high chair and sentences such as "the mouse is under the chair"? According to the classical view (e.g., Katz, 1972; Katz & Fodor, 1963), the concept chair represents (or consists of) information about the necessary and sufficient attributes of chairs. According to the family resemblance view (e.g., Rosch & Mervis, 1975), the concept chair is a summary representation that abstracts across specific instances of chairs to give information about what chairs, on average, are like. According to the exemplar view (e.g., Medin & Schaffer, 1978), the concept chair consists of representations of past exemplars of chairs that a person has experienced, rather than a single summary representation of all chairs. The schema view suggests that the concept chair consists both of representations of chair exemplars and of information about what chairs, on average, are like. Finally, according to the explanation-based view (e.g., Johnson-Laird, 1983; Lakoff, 1987b; Murphy & Medin, 1985), the concept chair includes information about the interaction among chairs, people, and other objects, as well as information about the (often, causal) relationships that hold among the different properties of chairs.

Many psychological studies up through the 1960s (e.g., Bourne, 1970; Bruner, Goodnow, & Austin, 1956; Hull, 1920) assumed that the classical view provided a proper description of most everyday lexical concepts. In the 1970s, evidence of pervasive incompatibilities with the classical view led to the view's decline. To replace the classical view and to account for these data, psychologists developed the family resemblance and exemplar views. Given the growing awareness at that time among psychologists of the philosophical work of Wittgenstein (1953), this seemed to be a step in the right direction. But by the mid-1980s and late 1980s, it became clear that the family resemblance and exemplar views (and simple hybrids of the two, such as some forms of the schema view) did not provide an adequate explanation of why certain classifications were privileged and tended to cohere. To deal with this problem, several cognitive scientists independently developed explanation-based views of concepts. (See Medin, 1989, for an excellent, brief overview of this chronology.)

In the first section of this article, after some introductory

This research was supported in part by a faculty development grant from the Ford Foundation administered by Carleton College.

I would like to thank Kathleen M. Galotti, John Macnamara, Laura Hipley, Therese Huston, Emilie Lin, Kenneth Thomases, and Amy Wilson for comments on a draft of this article.

Correspondence concerning this article should be addressed to Lloyd K. Komatsu, Department of Psychology, Carleton College, One North College Street, Northfield, Minnesota 55057-4025.

comments, I review the classical, family resemblance, and exemplar views. Although these three views were the subject of a classic review by E. E. Smith and Medin (1981), additional problems with the family resemblance view and additional support for the exemplar view have come to light since that time. I end the section with a discussion of the schema view. The relationship between the schema view and the family resemblance and exemplar views has not generally been discussed, but I argue that the former can be understood as a hybrid of the latter two, incorporating characteristics of both. I also argue that the schema view incorporates (but typically does not highlight) characteristics of the explanation view.

To provide a theoretical background for the explanation-based view, I begin the second section with a discussion of the theory of direct reference and its relevance for a theory of meaning as outlined by Putnam (1975a, 1975b). Putnam argues that meaning should not be equated with concepts and that concepts do not completely specify the objects to which a word refers. He suggests that the concepts associated with certain terms are much more like theories than they are like simple lists of attributes. I argue that some (but not all) of the issues raised by Putnam's analysis distinguish the explanation-based view from the similarity-based views. After a review of several variants of the explanation-based view, I conclude that an understanding of concepts can only emerge from an understanding of similarity, inference, and explanation—issues that highlight the ways in which we use and relate concepts to one another.

Some Preliminary Issues

Before reviewing the various views of concepts, a few preliminary issues need to be identified and clarified.

Processes and Representations

First, experimental results do not directly indicate anything about conceptual representations. It is always difficult to decide whether a particular observation is a function of the information represented by the concept, the structure, or the form of that information (e.g., propositional or imaginal; Anderson, 1978) or of the processes that operate on that concept (Armstrong, Gleitman, & Gleitman, 1983; Lakoff, 1987b; see E. E. Smith & Medin, 1981, for an extended discussion of the interaction between representation and process). Such results, however, do specify at a minimum what kind of information must be available to people about the instances of a concept (see Rosch, 1975b; Rosch & Mervis, 1975, for a similar argument). In general, recent views of concepts have focused on the contents of conceptual representations; that focus is maintained in this article.

Focusing on Everyday Objects

Second, because most recent views of concepts have been developed with particular reference to objects, discussion here is limited to concepts for everyday objects or classes of objects. The danger to this approach is that it is unclear whether a view developed through a consideration of objects will work for event concepts or other, more abstract concepts. The advantage to this approach is that it lessens the possibility that our theory building will be led astray by an attempt to capture a heterogeneous group. (Although it does not completely eliminate the problem, of course: An object category may be represented in more than one way, each implicated in a different task [e.g., classification, deduction, judgments of similarity].)

Coherence and Naturalness

A third issue that needs to be identified is that of coherence and naturalness. An adequate theory of concepts should explain what holds the extension of a word, the class of things to which it applies, together (and apart from things not in the class). *Coherence* refers to this holding together of instances. An adequate theory of concepts should also explain why certain classes seem to be "natural" and others do not. Why is it that the class of things that are either more than a second long or weigh less than a ton seems unnatural, but the class of things that have feathers and fly seems quite natural (Goodman, 1972; Osherson, 1978)? An understanding of coherence and naturalness would help us to understand why some categories, some ways of dividing up the world, are privileged over others.

Economy and Informativeness

The final preliminary issue is the trade-off between informativeness and economy. The world is categorized in part for reasons of economy. To remember and treat everything in one's environment as unique would require tremendous cognitive capacity (Anderson, 1991a). To say that two objects are both properly labeled by the same word, or are both instantiations of the same concept, is to say that for some purposes they can be treated as identical. Cognitive economy results to the extent that the characteristics that distinguish different instantiations of the concept can be ignored for such purposes.

But categorization is also done for reasons of informativeness. Categorization allows one to go beyond the information given. Once a person knows that a particular item is an instance of a particular concept, he or she can assume that it will be like other instances of the concept in certain ways. The greater the number and specificity of categories, the greater the homogeneity of their members, the greater the number of attributes those members reliably share, and the greater the informativeness of the categorization. The fewer or more general the categories, the lower their homogeneity and the lower their informativeness. Thus, economy and informativeness trade off against each other: If categories are very general, there will be relatively few categories (increasing economy), but there will be few characteristics that one can assume different members of a category share (decreasing informativeness) and few occasions on which members of the category can be treated as identical. If categories are very specific, there will be relatively many categories (decreasing economy), but there will be many characteristics that one can assume different members of a category share (increasing informativeness) and many occasions on which members can be treated as identical.

One of the crucial contributions made by Eleanor Rosch during the 1970s was the notion of a basic level of categorization (a notion that built on some comments by Brown, 1958). The basic level of categorization is the level of abstraction that represents the best compromise between number and informativeness of categories (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Chair is at the basic level of abstraction for most people in most situations. For most people in most situations, the superordinate category furniture is too generalized to be terribly useful. The subordinate category recliner, in contrast, typically does not add enough useful information to justify the greater number of distinctions it implies.

It is less often noted that the trade-off between economy and informativeness also arises within a given level of abstraction. To maximize cognitive economy, one can maintain that concepts are unitary: A single representation serves an entire category, representing all instances for all purposes. By this view, concepts are abstractions, recording the commonalities that cut across their different instances. Given that the instances of a category are not identical, this necessarily leads to a loss either in the amount of detail that a concept conveys or in the certainty with which one can assume that the details that are given are true of all instances. In contrast, to maximize informativeness, one can maintain that concepts are multiplex in nature: that a single word or mental category implies multiple representations. At the extreme, the concept associated with a word may be a collection of representations of individual instances. This would maximize informativeness (because a complete set of true information is retained for each instance) but would achieve no economy.

To a large extent, the different views of concepts reflect differing responses to the tension between informativeness and economy and to the problems of naturalness and coherence. These issues are recurring themes in this article.

Similarity-Based Views

Medin and his colleagues (Medin, 1989; Medin & Wattenmaker, 1987; Murphy & Medin, 1985) refer to the classical, family resemblance, and exemplar views as being similarity based. All three views suggest that objects are classified as instances of a category by virtue of the attributes they share with (i.e., their similarity to) some abstract specification of the category, or with known instances of the category. Beyond that common assumption, however, the three similarity-based views are quite different. Here I quickly review the similarity-based views, focusing on how each deals with the various thematic issues. (For further details on these models, see E. E. Smith & Medin, 1981.)

The Classical View: Necessary and Sufficient Characteristics

The distinguishing assumption of the classical view is that concepts are defined by sets of individually necessary and collectively sufficient attributes.[1] This has several important consequences: First, it implies that membership in a conceptual category is clear-cut (i.e., if an object has the necessary and sufficient characteristics, it is a member; if it does not, it is not). Second, it implies that membership in a conceptual category is discrete (i.e., an object is either a member or not, with no possibility of being more or less of a member).

Third, it implies a strong constraint on the attributes specified by a concept, namely, that they are individually necessary and collectively sufficient. Furthermore, it implies that such information (sometimes called *definitional* information; Katz, 1972; Miller, 1978) is all that is needed to explain how people understand linguistic relations (sometimes called *linguistic meaning;* Medin & Smith, 1984) and word combinations and how they make inferences about and recognize instances of the concept. Thus, the classical view implies a clear distinction between definitional information (i.e., information specified by a concept) and encyclopedic information (information about how the extension of the concept in the real world relates to other aspects of the world; see J. D. Fodor, 1977; Katz, 1972; Miller, 1978, for discussion).

Economy, Informativeness, Coherence, and Naturalness

The classical view leans heavily toward economy: A single representation is used for an entire category. It limits informativeness somewhat: Although the information in that representation can be assumed to be true of all instances, it is limited to just that information that will support a rather narrow range of conceptual phenomena (e.g., linguistic meaning). The classical view provides a strong account of category coherence: The members of a category are held together by the fact that they share a certain set of attributes with every other member and with no nonmembers (i.e., a set that is necessary and sufficient). The classical view does not address the issue of naturalness: Individual necessity and collective sufficiency are the only constraints on these attributes, so there is nothing to prevent the formation of arbitrary categories.

Decline of Necessity and Sufficiency

The assumption that concepts specify necessary and sufficient characteristics came under a great deal of attack during the 1970s (see Mervis & Rosch, 1981; Rosch, 1978; E. E. Smith & Medin, 1981, for reviews from that period). For example, when McCloskey and Glucksberg (1978) asked subjects to judge whether particular objects belonged to certain familiar categories, they found considerable disagreement among subjects (the modal response accounting for only 64% of the subjects in some cases) and considerable inconsistency within subjects (in some cases, subjects' answers across a 1-month period changed 22%

[1] At least since the 1950s, few psychologists have maintained that all concepts are described adequately by the classical view. For example, Bruner, Goodnow, and Austin (1956) discuss relational concepts (which are defined by a relation between two or more attributes), disjunctive concepts (which are defined by multiple sufficient sets, with few or no attributes being necessary) and even probablistic concepts. Effectively, therefore, there are no proponents of a pure classical view among psychologists. However, for ease of exposition, it is useful to assume such an idealized case; Bruner, Goodnow, and Austin, for example, clearly maintained that necessary and sufficient attribute sets play a key role in most concepts. The attacks on the classical view therefore can be seen as attacks on the usefulness of necessity and sufficiency in describing natural concepts.

of the time). This suggested that the boundaries between categories are fuzzy rather than clear-cut (see also Bellezza, 1984a; Hampton, 1979).

Also problematic for the classical view was the discovery that inspection of the definitions and attribute lists people generate for most everyday concepts contain considerable inter- and intrasubject variability (Bellezza, 1984b) and fail to meet the criteria of necessity and sufficiency (Barsalou, Spindler, Sewell, Ballato, & Gendel [cited in Barsalou, 1989]; Hampton, 1979, 1981), regardless of instructional manipulation (Ashcraft, 1978; Barsalou et al. [cited in Barsalou, 1989]; Komatsu, 1983; McNamara & Sternberg, 1983; Rosch & Mervis, 1975). (See J. A. Fodor, 1981, and J. A. Fodor, Garrett, Walker, & Parkes, 1980, for a different argument against definitions.)

Although fuzzy boundaries and the difficulty of obtaining definitions were noted quite early in the history of experimental research on concepts (cf. Smoke, 1932), such problems were usually regarded as explainable by a variety of factors. For example, some concepts (so-called *disjoint* concepts) may have multiple senses, each associated with a different set of necessary and sufficient attributes. A failure to specify exactly which sense of a word is intended would then lead to apparent boundary fuzziness and a difficulty in generating adequate definitions. The difficulty of giving definitions could also be explained by a lack of conscious access to definitions (rather than a lack of definitions per se) or by the fact that definitions typically are crucially dependent on attributes that are difficult to verbalize (e.g., perceptual attributes or "ineffable truth conditions," Johnson-Laird, 1987; McNamara & Miller, 1989).

More problematic for the classical view was evidence that categories are not discrete. Studies using a variety of methods demonstrated that subjects found it quite natural to judge different instances of many everyday concepts as being better or worse as examples (i.e., more or less typical) of those concepts (e.g., Berlin & Kay, 1969; Heider, 1972; Labov, 1973; Lakoff, 1972; Rosch, 1973, 1975b; Rosch & Mervis, 1975). More important, judged level of typicality (sometimes called *prototypicality*) was found to predict a variety of results. For example, level of typicality predicts reaction times in sentence verification tasks (Hampton, 1979; McCloskey & Glucksberg, 1979; Rips, Shoben, & Smith, 1973; Rosch 1973, 1975b; Rosch & Mervis, 1975; Rosch, Simpson, & Miller, 1976; E. E. Smith, Shoben, & Rips, 1974), order of item output when subjects are asked to name instances of a concept (Barsalou & Sewell, 1985; Mervis, Catlin, & Rosch, 1976), efficacy in priming tasks (Rosch 1975b; Rosch, Simpson, & Miller, 1976), and use as a cognitive reference point (compare the naturalness of "an ellipse is essentially a circle" to "a circle is essentially an ellipse"; Rosch 1975a; cf. also Lakoff, 1972).

The pervasiveness of typicality effects convinced at least some researchers (e.g., Rosch & Mervis, 1975) that they were the main phenomena to be explained. To the extent that the assumption of necessity and sufficiency does not allow the classical view to predict category gradedness, these typicality effects are an embarrassment for the view (E. E. Smith & Medin, 1981). Therefore, many researchers came to believe that the criteria of necessity and sufficiency, hence the classical view, had to be rejected. In its place, many accepted a position usually called the *family resemblance view.*

The Family Resemblance View: An Explanation for Typicality

According to the family resemblance view (Wittgenstein, 1953), categories cohere by virtue of the "family resemblances" among their members. Rosch and Mervis (1975), in one of the earliest and most influential studies to adapt Wittgenstein's philosophical views to psychology, describe family resemblances as follows:

> A family resemblance relationship consists of a set of items [in which] each item has at least one, and probably several, elements in common with one or more other items, but no, or few, elements are common to all items. . . . Members of a category come to be viewed as prototypical of the category as a whole in proportion to the extent to which they bear a family resemblance to (have attributes which overlap those of) other members of the category. Conversely, items viewed as most prototypical of one category will be those with least family resemblance to or membership in other categories. (p. 575)

Characterizing the Family Resemblance View

Five characteristics are typically associated with the family resemblance view. Several are apparent from the description quoted above. Others are not so apparent but follow in relatively straightforward fashion.

Centrality of typicality. First, as mentioned above, this description implies that degree of (proto)typicality, which is precisely what the classical view has difficulty explaining, is the central conceptual phenomenon. In the family resemblance view, to understand concepts, one must understand typicality. Degree of typicality is directly linked to degree of family resemblance: Items with greater family resemblance to a category are judged to be more typical of the category.

Abstractness. Second, every attribute specified for a concept is shared by more than one instance of the concept. Thus, the information contained in a concept is an abstraction across instances of the concept. The overlapping networks of shared attributes thus formed hold conceptual categories together. In this respect, the family resemblance view is like the classical view: Both maintain that the instances of a concept cohere because they are similar to one another by virtue of sharing certain attributes.

Weighted attributes. An object that shares attributes with many members of a category bears greater family resemblance to that category than an object that shares attributes with few members. This suggests that attributes that are shared by many members confer a greater degree of family resemblance than those that are shared by a few. A third characteristic of the family resemblance view is that it assumes that concept attributes are "weighted" according to their relevance for conferring family resemblance to the category.[2] In general, that relevance is taken to be a function of the number of category instances (and perhaps noninstances) that share the attribute.

Presumably, if the combined relevance weights of the attributes of some novel object exceed a certain level (what might be called the membership *threshold* or *criterion*), that object will be

---

[2] Here and throughout, I use *relevance* to include both *relevance* and *salience* as used by Ortony, Vondruska, Foss, and Jones (1985).

considered an instance of the category (Medin, 1983; Rosch & Mervis, 1975; E. E. Smith & Medin, 1981). The greater the degree to which the combined relevance weights exceed the threshold, the more typical an instance it is (see also Shafir, Smith, & Osherson, 1990). By this measure, an object must have a large number of heavily weighted attributes to be judged highly typical of a given category. Because such heavily weighted attributes are probably shared by many category instances and relatively few noninstances, an object highly typical of a category is likely to lie near the central tendencies of the category (see Retention of Central Tendencies, below), and is not likely to be typical of or lie near the central tendencies of any other category.

Independence and additive combination of weights: Linear separability. Attribute weights can be combined using a variety of methods (cf. Medin & Schaffer, 1978; Reed, 1972). In the method typically associated with the family resemblance view (adapted from Tversky's, 1977, contrast model of similarity), attribute weights are assumed to be independent and combined by adding (Rosch & Mervis, 1975; E. E. Smith & Medin, 1981). This leads to a fourth characteristic of the (modal) family resemblance view: It predicts that instances and noninstances of a concept can be perfectly partitioned by a linear discriminant function (i.e., if one was to plot a set of objects by the combined weights of their attributes, all instances would fall to one side of a line, and all noninstances would fall on the other side; Medin & Schaffer, 1978; Medin & Schwanenflugel, 1981; Nakamura, 1985; Wattenmaker, Dewey, Murphy, & Medin, 1986). Thus the (modal) family resemblance view predicts that concepts are "linearly separable."

Retention of central tendencies. The phrase *family resemblance* is used in two ways. In the sense that I have focused on until now, the family resemblance of an object to a category increases as the similarity between that object and all other members of the category increases and the similarity between that object and all nonmembers of the category decreases. This use of *family resemblance* (probably the use more reflective of Wittgenstein's, 1953, original ideas) has an extensional emphasis: It describes a relationship among objects and makes no assumptions about how the category of objects is represented mentally (i.e., about the intension of the word or what I have been calling the *concept*).

In the second sense, family resemblance increases as the similarity between an object and the central tendencies of the category increases (Hampton, 1979). This use of *family resemblance* has an intentional emphasis: It describes a relationship between objects and a mental representation (of the central tendencies of a category).

Although these two ways of thinking about family resemblance, average similarity to all instances and similarity to a central tendency, are different (cf. Reed, 1972), Barsalou (1985, 1987) points out that they typically yield roughly the same outcome, much as the average difference between a number and a set of other numbers is roughly the same as the difference between that number and the average of that set of other numbers. (For example, consider the number 2 and the set of numbers 3, 5, and 8. The average difference between 2 and 3, 5, and 8 is 3.33, and the difference between 2 and the average of 3, 5, and 8 is 3.33.) Barsalou argues that although for most purposes the two ways of thinking about family resemblance are equivalent (one of the reasons the exemplar and family resemblance views are often difficult to distinguish empirically; see below), computation in terms of central tendencies may be more plausible psychologically (because fewer comparisons are involved in comparing an object with the central tendencies of a concept than with every instance and noninstance of the concept; see also Barresi, Robbins, & Shain, 1975). This suggests a fifth characteristic of the family resemblance view: A concept provides a summary of a category in terms of the central tendencies of the members of that category rather than in terms of the representations of individual instances.[3]

Economy, Informativeness, Coherence, and Naturalness

Both the classical and the family resemblance views explain conceptual coherence in terms of the attributes shared by the members of a category (i.e., the similarity among the instances of a concept). The critical difference between the two views lies in the constraints placed on the attributes shared. In the classical view, all instances are similar in that they share a set of necessary and sufficient attributes (i.e., the definition). The family resemblance view relaxes this constraint and requires only that every attribute specified by the concept be shared by more than one instance.

Although this requirement confers a certain amount of economy to the family resemblance view (every piece of information applies to several instances), removing the definitional constraint allows family resemblance representations to include nondefinitional information. In particular, concepts are likely to specify information beyond that true of all instances or beyond that strictly needed to understand what Medin and Smith (1984) call *linguistic meaning* (the different kinds of relations that hold among words such as *synonymy, antynomy, hyponomy, anomaly,* and *contradiction* as usually understood; cf. Katz, 1972; Katz & Fodor, 1963) to include information about how the objects referred to may relate to one another and to the world.

It is not clear whether this loss in economy results in a concomitant increase in informativeness: Although in the family resemblance view more information may be associated with a concept than in the classical, not all of that information applies to every instance of the concept. In the family resemblance view, attributes can be inferred to inhere in different instances only with some level of probability. Thus the informativeness of the individual attributes specified is somewhat compromised.

With no a priori constraint on the nature (or level) of similar-

[3] There are several different ways to approach the representation of the central tendencies of a category. E. E. Smith and Medin (1981), for example, identified three approaches to what they called the probabilistic view: the featural, the dimensional, and the holistic. E. E. Smith and Medin provided ample evidence for rejecting the holistic approach on both empirical and theoretical grounds (see also McNamara & Miller, 1989). They also argued that the similarities between the featural and dimensional approaches suggest that they might profitably be combined into a single position that could be called the "component" approach (E. E. Smith & Medin, 1981, p. 164) and concluded that the component approach is the only viable variant.

ity shared by the instances of a concept, the family resemblance view has difficulty specifying which similarities count and which do not when it comes to setting the boundaries between concepts. A Great Dane and a Bedlington terrier appear to share few similarities, but they share enough so that both are dogs. But a Bedlington terrier seems to share as many similarities with a lamb as it does with a Great Dane. Why is a Bedlington terrier a dog and not a lamb?

Presumably, the family resemblance view would predict that the summed weights of Bedlington terrier attributes lead to its being more similar to other dogs than to lambs and result in its being categorized as a dog rather than a lamb. But to determine those weights, we need to know how common those attributes are among dogs and lambs. This implies that the categorization of Bedlington terriers must be preceded by the partitioning of the world into dog and lamb. Without that prior partitioning, the dog versus lamb weights of Bedlington terrier attributes cannot be determined. To answer the question of what privileges the categorization of a Bedlington terrier with the Great Dane rather than the lamb requires answering what privileges the partitioning of the world into dogs and lambs.

Rosch (Rosch, 1978; Rosch & Mervis, 1975) argues that certain partitionings of the world (including, presumably, into dogs and lambs) are privileged, more immediate or direct, and arise naturally from the interaction of our perceptual apparatus and the environment. Thus whereas the classical view stresses the coherence of conceptual categories without addressing naturalness, the family resemblance view stresses naturalness with coherence emerging as a by-product (Neisser, 1987). Whereas the classical view constrains concepts through an abstract specification of the attributes that constitute a concept (i.e., that they are individually necessary and collectively sufficient), the family resemblance view suggests that concepts are constrained (at least at the basic level of abstraction) ecologically, reflecting the natural partitioning of objects in the real world by our perceptual systems. (See Anderson, 1990, 1991a, 1991b for a closely related, and more fully developed, view of human categorization as adaptive.)

A second constraint on concepts operates in the (modal) family resemblance view: linear separability. The assumption that weights are combined by summing allows the family resemblance view to describe only those categories that linearly partition the attribute space. In some sense, therefore, this view claims that category boundaries are linear discriminant functions (Murphy & Medin, 1985); items cohere by virtue of falling on the same side of some such function. Notice that unlike the ecological constraint, linear separability addresses coherence but does not address naturalness (because any arbitrary set of attributes may be used to define the attribute space). But presumably, the ecological constraint may favor certain discriminant functions over others. Thus the ecological and linear separability constraints may work together to allow the family resemblance view to explain why some categories are privileged over others.

Ascendence of the Family Resemblance View

Although there is a certain degree of vagueness in the family resemblance view,[4] it is clear that it holds a great deal of promise for explaining a variety of phenomena that are difficult for the classical view to accommodate. For example, because the family resemblance view rejects the notion that the attributes specified by a concept are necessary and sufficient, it explains the failure of subjects to give necessary and sufficient definitions and attributes in a very straightforward manner: There are none to be given.

Explaining typicality effects is another strong suit of the family resemblance view. In fact, E. E. Smith and Medin (1981, p. 69) argue that typicality effects follow so naturally from the family resemblance view that such effects can be considered to be support for this view. Furthermore, because the family resemblance view explains both typicality and category membership in terms of the values obtained by combining attribute relevance weights, it has a natural means of explaining fuzzy boundaries between instances and noninstances: Any noise or variability in the relevance weights of any attribute would lead to a fuzzy boundary.

Because of the wealth of empirical data demonstrating the pervasive fuzziness of concepts and the wide range of results predicted by typicality judgments, the family resemblance view rapidly gained acceptance among cognitive psychologists (cf. textbooks by Bourne, Dominowski, & Loftus, 1979; Glass, Holyoak, & Santa, 1979; Klatzky, 1980). Soon, family resemblance analyses (usually called *prototype analyses*) were being applied not only to object concepts but also to emotions (Fehr, 1988; Fehr & Russell, 1984; Shaver, Schwartz, Kirson, & O'Connor, 1987), trait and person concepts (Cantor & Mischel, 1977; Mayer & Bower, 1986), psychological situations (Cantor, Mischel, & Schwartz, 1982; see Lingle, Altom, & Medin, 1984, for general discussion about the application of prototype analyses to social psychological categories), and clinical categories or categories of abnormal behavior (Cantor, Smith, French, & Mezzich, 1980; Genero & Cantor, 1987; Horowitz, Wright, Lowenstein, & Parad, 1981), as well as to styles of painting (Hartley & Homa, 1981) and musical themes (Welker, 1982).

Dissatisfaction With the Family Resemblance View

Unfortunately, it soon became clear that rejecting the classical constraint of necessity and sufficiency led the family resemblance view into some difficulty. For example, how are complex concepts (e.g., pet fish or "the center dunked the basketball") constructed out of simple(r) concepts (e.g., pet and fish)? The classical account built on set theory and described complex concepts as the union of the necessary and sufficient attribute sets of the constituent simple concepts. Because sets of necessary and sufficient attributes do not exist in the family resemblance view, that account cannot be directly adopted.

---

[4] For example, one can assume in this view either that conceptual representations include information about the central tendencies of all attributes displayed by the instances of a category or of only some (unspecified) subset of attributes (typically, those that are most relevant or heavily weighted). The former position (which might be called the *prototype* approach) argues that concepts represent potential instantiations, whereas the latter (which might be called the *cluster* approach) does not (see also Kelley & Krueger, 1984; Reed, 1972; E. E. Smith & Medin, 1981).

Initially, supporters of the family resemblance view appealed to logics that were based on fuzzy sets. The idea was that the extension of most everyday terms were fuzzy sets (Hersh & Caramazza, 1976; Rosch & Mervis, 1975) and fuzzy-set logic would therefore predict the properties of combinations of such terms. Unfortunately, such accounts were found to be inadequate (Cohen & Murphy, 1984; Johnson-Laird, 1983; Osherson & Smith, 1981, 1982; but see Zadeh, 1982). Later work suggested that a family resemblance approach focusing on the subjective weightings of attributes that characterize instances of a concept may be able to deal with conceptual combination without appealing to fuzzy-set theory (e.g., Hampton, 1987, 1988; Shafir et al., 1990; E. E. Smith, 1988; E. E. Smith & Osherson, 1984; E. E. Smith, Osherson, Rips, & Keane, 1988; but see Medin & Shoben, 1988; Murphy, 1988).

A second strength of the classical view lost by the family resemblance view is in accounting for linguistic meaning (Katz, 1972; Katz & Fodor, 1963; Medin & Smith, 1984). Although some progress has been made recently in explicating inductive reasoning within the family resemblance view (Osherson, Smith, Wilkie, López, & Shafir, 1990; Rips, 1975), the family resemblance view still finds it difficult to account for certain forms of linguistic relations.[5]

A third problem for the family resemblance view is that people (or at least people in literate societies) have strong intuitions that words have necessary and sufficient definitions, despite the fact that they cannot articulate those definitions (cf. McNamara & Sternberg, 1983). This intuition, in fact, may be the main reason that the classical view held sway for so long. People also seem to have the intuition that the boundaries for most everyday categories are clear-cut, although those intuitions may not be terribly pervasive or strong (cf. Armstrong et al., 1983).

Fourth, explaining naturalness and coherence in terms of the interaction between the human perceptual system and the environment means that the family resemblance view accounts for these characteristics only in perceptually based concepts (e.g., the domain of color). But even as he argues for the importance of an ecological constraint, Neisser (1987) notes that a perception-based account of concepts is ultimately inadequate; few adult concepts rely exclusively on perceptual similarity.

Finally, subsequent research has indicated that typicality effects in themselves do not clearly dictate the nature of conceptual structure. In particular, they do not provide good justification for rejecting necessity and sufficiency. For example, Armstrong et al. (1983; Gleitman et al., 1983; see also Bourne, 1982) found that typicality judgments can be obtained for concepts with clear definitions (e.g., odd number) and that such judgments will predict reaction times in the usual tasks. Although typicality effects may indicate that the classical view is not adequate for capturing all conceptual phenomena, they do not rule out the possibility that the classical view may describe at least some aspects of many concepts.

A possible fix: Combining the classical and family resemblance views. The various inadequacies of the family resemblance view led a number of psychologists (e.g., Landau, 1982; Miller & Johnson-Laird, 1976; Neimark, 1983; Rosch, 1983; E. E. Smith, 1988; E. E. Smith & Medin, 1981) to propose hybrid or dual-representational models that include both classical and family resemblance representations.

Most of the details of how classical and family resemblance representations would divide and coordinate the labor of accounting for different conceptual phenomena have not been worked out. In most such cases (Landau, 1982; Miller, 1978; Miller & Johnson-Laird, 1976; Osherson & Smith, 1981; E. E. Smith, 1988; E. E. Smith & Medin, 1981), the family resemblance representation is argued to be the basis for identifying instances of the concept, whereas the classical representation is the basis for reasoning about concepts (sometimes called the core/identification procedure approach). In others, the two representations map onto a distinction between competence and performance (Neimark, 1983) or between logical and reference point (in effect, a kind of analogical) reasoning (Rosch, 1983).

For a dual-representational approach to work, each aspect must hold up its own end of the explanatory burden. But the family resemblance view appears to be unable to account for several results that such hybrids assume it explains. These failures stem primarily from two built-in limitations of the view:

Limiting concepts to those that are linearly separable. Independence of attribute weights and combination of weights through summing are usually assumed in the family resemblance view. These assumptions mean that the family resemblance view only describes linearly separable concepts. This implies that linearly separable artificial categories are somewhat more natural and should be easier to learn than nonlinearly separable ones. But the available evidence (Kemler Nelson, 1984; Medin & Schwanenflugel, 1981; Nakamura, 1985) does not support this prediction. In fact, under certain instructional conditions, linearly separable categories may be more difficult to learn (Wattenmaker, Nakamura, & Medin, 1988).

An easy fix for this problem would be to surrender the assumption of additive combinations of attribute weights. Using alternative combination methods (e.g., multiplicative; cf. Medin & Schaffer, 1978) would allow the family resemblance view to account for both linearly and nonlinearly separable categories. Because this assumption is not critical to the characterization

[5] However, this problem may be more apparent than real. It is possible that at least some of the phenomena of linguistic meaning that are most problematic for the family resemblance view (i.e., those that assume logical entailment) simply do not exist. Quine (1953), for example, argues that so-called *analytic statements* can be revised and are not true (as usually described) by definition. Such statements resist revision only because their revision would require the revision of many other statements. For example, it is not necessarily true that given a line and a point not on that line, it is possible to draw one and only one line through the point that is parallel to the first. But to revise that statement (i.e., to revise the notion of parallel lines) would require revising many other statements about geometry. Baker (1974), echoing Wittgenstein (1953), argued that the notion of criteria should be substituted for logical necessity. The criterion relationship is one that is established by convention. It is therefore somewhat stronger (i.e., more definite) than a simple empirical generalization that is based on probabilities (i.e., an induction) but not as definite as the relationship of logical entailment conveyed by necessity (Johnson-Laird, 1983). Therefore, although linguistic meaning poses a problem for the family resemblance view (and, as it turns out, the exemplar view) for now, future work on the notions of criteria and differential revisability may provide an adequate account of what have traditionally been interpreted as the empirical consequences of logical entailment.

of the family resemblance view, such a modification seems acceptable.

Limiting concepts to information about central tendencies and attribute weights. More critical to the family resemblance view is the assumption that a concept only represents information about the central tendencies and relative weights of the attributes that characterize the instances of the category. But subjects who learn a category also seem to have information about the variability of the instances of the category (Barresi et al., 1975; Homa & Vosburgh, 1976) and about the correlations among the attributes of the instances in the category (Malt & Smith, 1984; Medin, Altom, Edelson, & Freko, 1982). Similarly problematic for the family resemblance view are the effects expectations regarding the distributions of instances have on category formation (Flannagan, Fried, & Holyoak, 1986) and the fact that subjects learn about categories in a sorting task when given no feedback about the correctness of their sorts (Fried & Holyoak, 1984; but see Homa, Burruel, & Field, 1987; Homa & Cultice, 1984, for other results and interpretations).

There is also good evidence that similarity to central tendencies does not fully explain typicality effects for all categories. In the case of ad hoc, goal-derived categories (e.g., "things to take from one's home during a fire"; Barsalou, 1983, 1985) or certain abstract categories (e.g., a belief or an instinct; Hampton, 1981), similarity to central tendencies seems to play very little role in determining typicality. In such cases, similarity to an ideal or frequency of instantiation seems to be a more powerful determinant (Barsalou, 1985). In some cases in which similarity to central tendency clearly does play a role (e.g., birds), similarity to an ideal and frequency of instantiation may contribute to typicality as well (Barsalou, 1985; Nosofsky, 1988b).

Finally, although the fuzziness of concept boundaries (McCloskey & Glucksberg, 1978) may explain some inconsistencies in typicality judgments both within and between subjects (Barsalou, 1987; Barsalou & Sewell, 1984 [cited in Barsalou, 1989]), it cannot explain why typicality judgments vary systematically with context (Barsalou, 1985; Roth & Shoben, 1983) and points of view (Barsalou, 1987). In general, because the representations described by the family resemblance view are context free, they cannot explain how levels of family resemblance or relevance weights of attributes (hence typicality judgments) are affected by context.

A possible fix: Multiple family resemblance representations. These problems might be explained by proposing multiple representations for each concept in a fashion somewhat akin to disjunctive concepts in the classical view. Instead of bird being linked to a single representation, there may be one (family resemblance) representation for birds of prey, a second for song birds, a third for fowl, and so on. Instability and sensitivity to context or point of view could then be explained by different subjects' calling up different bird representations under different circumstances. Such a multiple-representation approach is sometimes assumed implicitly in family resemblance discussions (e.g., Rosch, 1978).

But consider the following: Suppose one wishes to determine whether a particular object is an instance of a bird. According to the family resemblance view, to do that one must determine the birdness-establishing relevance weights of the object's attributes. Suppose the object flies. The multiple-representation approach does not give a single weight for flying. It gives different weights depending on the particular representation (bird of prey, fowl, and so on) accessed. With a multiple-representation account of the concept bird, what is really decided is whether the object is an instance of bird of prey or fowl, not whether it is a bird as such. To decide whether an object is a bird takes two steps: First, determine whether the object is an instance of bird of prey or fowl, and if it is, then, second, infer that it is also an instance of (the superordinate category) bird.

The crucial point here is that the multiple-representation account of bird does not provide a family-resemblance-based account of bird, but only one for specific types of birds. A multiple-representation account of bird is better understood as an example of the exemplar rather than the family resemblance view (E. E. Smith & Medin, 1981).

The Exemplar View

In the classical view, every instance has every attribute specified by the conceptual representation (although this is somewhat relaxed in allowing for disjunctive concepts). In the family resemblance view, it is rarely the case that a given attribute specified by the representation will characterize every instance. But because attributes unique to specific instances of a concept do not contribute to family resemblance, unique attributes are not included in a family resemblance representation. Therefore, every attribute specified by a family resemblance representation characterizes more than one instance. In the exemplar view, even this latter requirement is relaxed: Attributes specified for a concept need not hold true for more than one instance. Thus, a concept ends up being a set of representations, with individual representations corresponding to a different exemplar of that concept. Unfortunately, it is not entirely clear what an exemplar representation is.

At one extreme, an exemplar representation may be a family resemblance representation that abstracts across different specific instances. The case of the concept bird discussed above is an example of this approach: The concept bird consists of the set of family resemblance representations corresponding to canary, robin, pigeon, duck, penguin, and so on. I call this the *multiple-prototype* approach.

At the other extreme, exemplar representations may involve no abstraction. Each of the representations making up the concept is a memory trace of a previously encountered instance. I call this the *instance* approach.

There has been little discussion of the multiple-prototype approach (although E. E. Smith & Medin, 1981 suggest that this is a possible interpretation of Medin & Schaffer, 1978, and it is clearly related to the modified multiple-prototype model Lakoff, 1987a, 1987b, proposes; see also Anderson, 1990). Instead, most attention has been focused on whether abstraction is a *necessary* part of conceptual representation (i.e., the adequacy of the instance approach). Therefore, my discussion of the exemplar view is limited to the instance approach.

Characterizing the Instance Approach

Similarity-based judgments. The exemplar view, like the classical and family resemblance views, is still fundamentally a similarity-based view. An item is judged to be an instance of a concept to the extent that it is sufficiently similar in the relevant fashion to one or more of the instance representations that constitute the concept. Concepts cohere because their constitutive instances are similar to one another in particular ways.

Abstracting across instances. The instance approach does not assume that abstraction across instances never takes place. It does assume, however, that if abstraction across instances takes place, it takes place when using the concept (e.g., making a typicality judgment, deciding whether a novel object is an instance, reasoning about the concept) rather than when learning it. In contrast, the family resemblance view assumes that abstracted information is stored as part of the concept and preexists an attempt to use the concept. (See Medin & Smith, 1981; E. E. Smith, 1978, for further discussion of the distinction between prestored information and information computed online.) According to the family resemblance view, then, abstraction across instances takes place during acquisition (or shortly thereafter).

Variants of the instance approach. It is possible to make different assumptions about the number and nature of instances stored. For example, it is possible to believe that every instance encountered is stored (as in the proximity model, described by Reed, 1972) or that only the best, most typical, or most frequent instances are stored (e.g., the best-examples model, which E. E. Smith & Medin, 1981, suggest is implicit in Rosch's, 1975a, discussion of cognitive reference points), or (as is assumed by most who take the instance approach) that most or many encountered instances are stored, to varying degrees of completeness.[6] Different variants of the instance approach are described by Brooks (1978), Hintzman (1986), Hintzman and Ludlam (1980), Medin (1986), Medin and Schaffer (1978), Nosofsky (1984, 1986, 1988a, 1988b, 1991), and Whittlesea (1987).

Choosing Between Family Resemblance and Exemplar Views: Empirical Considerations

By suggesting that representations of one or more instances are retrieved from long-term memory and compared with the object at hand or used as the basis for performing the task at hand, the instance approach can explain many of the same results as the family resemblance view. Because different representations may be accessed at different times, instance models have a very natural (if perhaps unsatisfyingly vague at this point) way to account for polysemy (cf. Palmer, 1981; Panman, 1982), fuzzy boundaries, typicality effects with sentence verification and priming tasks, and people's inability to give definitions.

Results favoring the exemplar view. Although the exemplar view, like the family resemblance view, has difficulty accounting for linguistic meaning, the instance approach can account for other results that are problematic for the family resemblance view. For example, by assuming that the particular instances retrieved from long-term memory depend on context, goals, prior processing, frequency of occurrence or retrieval, time of last retrieval, and so on, the instance approach can explain specificity of encoding (Brooks, 1987), the instability of instance retrieval and typicality judgments, and the sensitivity of typicality judgments to context, goals, points of view, and frequency of instantiation (Barsalou, 1987, 1989).[7] By assuming that expectations affect the particular instances (or attributes of those instances) that are retained, the instance approach can explain the effects of expectations on concept acquisition. The instance approach is also consistent with the finding that accuracy of classification increases with increases in category size (Busemeyer, Dewey, & Medin, 1984; Hintzman, 1986; but see Homa, Dunbar, & Nohre, 1991). By suggesting that the representations of more than one instance can be simultaneously retrieved and compared, the instance approach can explain subjects' sensitivity to correlations among the attributes of the instances of the category (Malt & Smith, 1984; E. E. Smith & Medin, 1981). If one assumes that subjects only retrieve a subset of their stored instances on any particular occasion but are inclined to regard that partial retrieval as exhaustive (see Nickerson, 1981), the instance approach may be able to explain beliefs in definitions and clear boundaries. Finally, by assuming a multiplicative rather than an additive combination of matching and mismatching attributes appropriately weighted (as in the context model; cf. Medin & Schaffer, 1978; Nosofsky, 1986), instance models provide an account for both linearly and nonlinearly separable categories (see also Medin & Schwanenflugel, 1981; Nakamura, 1985; Wattenmaker et al., 1986).

Problematic results: Availability of information about central tendencies. Experiments by Posner and Keele (1968) and Franks and Bransford (1971) showed that subjects in concept acquisition tasks seemed to have knowledge about the central tendencies (specifically, the prototypes) of a category, even if no instances presented to them exactly matched those central tendencies. Although these results were originally interpreted as support for some version of the family resemblance view, the instance approach easily explains them by arguing that subjects in these experiments computed the central tendencies during the test phase by retrieving representations of old instances. But there is a difficulty with this explanation: Information about central tendency remains available even after information about old instances has faded (Homa, Cross, Cornell, Goldman, & Shwartz, 1973; Homa & Vosburgh, 1976; Posner & Keele, 1970; Robbins et al., 1978; Strange, Keeney, Kessel, & Jenkins, 1970). Such results, on the other hand, are clearly consistent with the family resemblance view, which maintains that central tendencies are computed during the learning phrase.

Subsequent work (e.g., Busemeyer et al., 1984; Hintzman, 1986; Hintzman & Ludlam, 1980; Medin & Schaffer, 1978; No-

[6] Although it is not be an issue here, note also that for a given set of representational assumptions, one can make different processing assumptions (e.g., whether all or only some stored instances are retrieved whenever the concept is used and, if only some are retrieved, whether retrieval is purely stochastic or guided in some fashion).

[7] Barsalou (1987) argues that all conceptual effects depend on representations constructed in working memory. In many ways, this position is consistent with the exemplar view. However, whereas the exemplar view focuses on the nature of the information stored in long-term memory, Barsalou's (1987) view focuses on the processes that affect the retrieval of information from long-term memory and its assembly in working memory. As such, I am more inclined to regard Barsalou's (1987) view as an explanation-based theory.

sofsky, 1988a), however, has shown that these results may not pose a problem for the exemplar view after all. They can be predicted by an instance model if it is assumed that incomplete or biased memory traces for more than one instance are retrieved during the test phase.

Problematic results: Dissociation of classification and recognition. In an instance model, items are classified by comparing them to previously encountered instances. Thus, instance models appear to predict that classification of new instances depends on the ability to recognize previously presented items. But Hayes-Roth and Hayes-Roth (1977) showed that the link between recognition and classification often breaks down. Using a paradigm similar to that of Posner and Keele (1968), Hayes-Roth and Hayes-Roth (like Posner and Keele, 1968) found that subjects were more confident about their classification of the prototypes than of old instances, suggesting the storage of central tendency information. But the Hayes-Roth and Hayes-Roth subjects were also more confident that they had previously seen certain old instances than they were that they had previously seen the prototypes (suggesting the storage of instance information). Because classification and recognition were not perfectly correlated, the Hayes-Roth and Hayes-Roth results suggest that human concepts include both abstracted and particularized information (Anderson, Kline, & Beasley, 1979).

Similarly, Metcalfe and Fisher (1986) argue that the instance model should predict that subjects are better at classifying items that they recognize as being old (i.e., presented previously) than those they judge to be new (i.e., not previously presented). Although they did find that the conditional probability of correctly classifying an item given that it was recognized as being old was slightly greater than the probability of correctly classifying an item given that it was recognized as being new, the difference was not significant. They concluded that this result is also problematic for an instance approach that does not prestore any abstracted information.

Nosofsky (1988a; see also Medin, 1986), however, argues that these demonstrations of a dissociation between recognition and classification do not provide compelling reason to abandon the instance approach. He points out that an instance approach (specifically, an extension of the context model that assumes that multiple incomplete or biased representations of instances are retrieved) can predict each of these results (Ashby & Lee, 1991).

Economy, Informativeness, Coherence, and Naturalness

The greater informational richness of concepts described by the exemplar view is responsible for that view's superiority over the family resemblance view in accounting for certain results (e.g., sensitivity to attribute correlations). That informational richness is purchased at the cost of economy, however. At the extreme, if all information about every encountered instance were to be stored as part of the concept, no cognitive economy at all would be realized.

But the instance approach does not require that all information about every instance be retained. In fact, the success of most instance models in predicting nonintuitive outcomes (e.g., Hintzman, 1986; Hintzman & Ludlam, 1980; Nosofsky, 1988a, 1989, 1991) depends on the fact that certain instances or specific details of certain instances will be forgotten over time (or not encoded initially). The problem is that the instance models do not provide a systematic means for determining which instances or details are retained.

Furthermore, the instance approach gives no systematic explanation of what makes a conceptual category cohere. Although new instances of a category are identified by their similarity to known instances, no constraints whatsoever are placed on the particular objects that can be (or are more likely to be) placed, initially, in the category. This arbitrariness in some ways recalls the classical view, in which no initial constraints are placed on the particular attributes that may constitute the necessary and sufficient set that defines a category. The critical difference is that in the classical view, once that set of attributes is set up, there is a very strict constraint on possible new instances of the concept (namely, every new instance must have those definitional attributes). In contrast, the instance approach places no strong constraints on possible new instances. It simply specifies that new instances must be similar to one or several of the previously identified instances in some way. In contrast to the family resemblance view, which identifies new instances of a category by their similarity to the central tendency of the category, the instance approach allows for categories that are based entirely on so-called *edge matching* (e.g., one instance of the category may be similar to a second instance, which is similar to a third, but there may be little similarity between the first and the third instances). With no prior specification of the nature or degree of similarity necessary for items to be instances of the same concept, there is no constraint at all on possible new instances: At the extreme, every object is similar to every other object in some way (e.g., having mass, having a location in space; Goodman, 1972).

For the instance approach to explain what makes concepts cohere and what privileges the classifications that one makes, it must specify some constraints, either on the initial set of objects that are classified together or on the allowable new instances, or both. Adopting the ecological constraints (Neisser, 1987; Rosch, 1978) discussed in connection with the family resemblance view is one useful possibility in this regard.

A second possibility (not incompatible with the first) is to combine the family resemblance view and instance approach into an *exemplar hybrid.* Such an exemplar hybrid would propose that concepts store information both about specific instances and about central tendencies, much as described by Homa et al. (1991). Assuming that the two kinds of information are well integrated, such a hybrid would have certain functional advantages. Family-resemblance-type information about central tendencies could guide the selection and retention of new instances: Instances that do not significantly deviate from the central tendencies need not be retained (although, of course, they may be). The abstract family resemblance representation of the central tendencies of the category may provide the concept with whatever economy, coherence, and naturalness it might have.[8] On the other hand, the representation of particu-

[8] Note that the advantages I identify for an exemplar hybrid over an instance model do not include the ability to account for other kinds of data. Hintzman (1986) and Nosofsky (1988a) argue that if the full range

lar instances allows the exemplar hybrid to encode information about the correlations among attributes, variances in attributes, relationships among attributes, and so on. In the next section, I consider in detail a particular exemplar hybrid: the schema view.

The Schema View: An Exemplar Hybrid

The notion of a schema or schemalike representations, originally introduced to modern psychology by Bartlett (1932) and Piaget (1926), resurfaced under a variety of labels (e.g., *schema, frame, script*) in the mid-1970s, largely through the work of cognitive scientists concerned about the representation of knowledge in computers (e.g., Minsky, 1975; Rumelhart, 1980; Rumelhart & Ortony, 1977; Schank & Abelson, 1977; Winograd, 1975; an influential exception is Neisser, 1975, who wrote about the role of schemata in human attention and perception). Since that time, although the schema view has been criticized on the grounds of being too vague (cf. Fiske & Linville, 1980), schema-based explanations of a wide variety of phenomena have been proposed, particularly in the area of social psychology (e.g., Taylor & Crocker, 1980), text processing (e.g., Kintsch & van Dijk, 1978), and reasoning (e.g., Nisbett & Ross, 1980; E. E. Smith, 1989).

The criticism of vagueness is fair, but I think it stems more from a too-liberal use of the term *schema* than from any problem inherent with the schema view. In some work (e.g., Posner & Keele, 1968), the terms *schema* and *prototype* are used interchangeably. In other work (e.g., Anderson, 1980), *schema* and *prototype* are clearly distinguished. Others have used *schema* very generally, not suggesting any more than that schemata are networks of associations that include representations of specific instances (e.g., Bem, 1981). Here, I intend a schema to be a single structure that captures characteristics of both the family resemblance view (by storing information that is abstracted across instances) and the instance approach (by retaining information about actual instances). It provides a uniform method of simultaneously representing information at different levels of abstraction. The notion of a schema as described by Rumelhart (1980; Rumelhart & Ortony, 1977) and developed by Cohen and Murphy (1984) fits this description.

Characterizing Schemata

An early description of schemata was given by Rumelhart (1980):

> According to schema theories, all knowledge is packaged into units. These units are the schemata. Embedded in these packets of knowledge is, in addition to the knowledge itself, information about how this knowledge is to be used.
>
> A schema, then, is a data structure for representing the generic concepts stored in memory. There are schemata representing our knowledge about all concepts. . . . A schema contains, as part of its specification, the network of interrelations that is believed to normally hold among the constituents of the concept in question. (p. 34)

Although very general, this description highlights two kinds of information included in schemata that are rarely discussed (although typically not specifically excluded; cf. Kellogg, 1981) by the classical, family resemblance, and exemplar views: information about how conceptual information is to be manipulated and information about the relationships that hold among the attributes of the instances of a concept. The schema view has other characteristics that are not typically discussed in connection with the family resemblance or exemplar views or that are described using different terminology.

Slots and slot values. The information that constitutes a schema is described (and organized) in terms of variables called *roles* or *slots* (closely akin to what I have been calling *attributes*). For example, piano would include slots for size, color, function, sound, and location (Cohen & Murphy, 1984), as well as slots for constituent parts (e.g., legs, keys). The schema specifies the values that can and cannot fill each slot (e.g., the size slot for piano cannot be filled by blue) and may specify the probability distribution of values that the slot may be filled with (e.g., a piano is most likely to be of a size so that an adult can comfortably sit at it and still reach all its keys, although smaller and larger sizes are possible).

Default values. If no value for a slot is specified for a particular instance of a concept (e.g., if the color of a particular piano is not specified), a value for that slot, called the *default value,* is inferred. Several aspects of default values should be noted. First, default values can be overridden if the context on a particular occasion of use suggests a different value. For example, although one might ordinarily assume that pianos are purely acoustic instruments (the default value), that value can be overridden if one is presented with an electric piano that has built-in electronic amplification. Second, assignment of default values can be context free or contingent on the values assigned to other slots (Rumelhart & Ortony, 1977). For example, values for a given slot can be ordered by frequency of occurrence or central tendencies among instances so that slot values default in context-free fashion to the most frequent (or average) value (e.g., about 4.5 feet tall).

Relationships among slots: Schemata as networks. Recall that in addition to retaining information about slots and slot values, schemata are assumed to maintain information about the relationships among slots and slot values (Cohen & Murphy, 1984). The explicit coding of relationships among slots shows the kinship between the schema view and network models of semantic memory (cf. E. E. Smith, 1978). In fact, Minsky (1975, p. 212) suggests that frames (roughly, a method for notating schemata) are networks of nodes (roughly, slots) and relations between them.

Regarding concepts as networks calls attention to the fact

---

of data on concept usage can be explained by a model that assumes that only information about instances is prestored, then parsimony argues for accepting the instance approach. On the other hand, Schank, Collins, and Hunter (1986) argue that if people must be granted the capacity to abstract across instances, the instance approach is unparsimonious because it assumes, without compelling reason, that the ability to abstract is not applied in a particular situation (i.e., during acquisition). Oden (1987) similarly argues that it is more parsimonious to assume that at least some abstraction takes place at the time of learning.

that schemata actually include information about two kinds of relationships: those that hold among the constituent attributes of a concept (e.g., that the wings of a bird are placed on the upper sides of its body) and those that hold among concepts, such as class inclusion (e.g., that canaries are a type of bird). Explicitly noting class-inclusion relationships provides a straightforward route through which slots and slot values can be "inherited" as default values by "descendent" concepts. For example, the value flies may well be part of the schema for bird. In most cases, it may be assumed that subtypes of bird will inherit that value (i.e., the default assumption is that different types of birds will fly), but the value may be overridden in some cases (e.g., with penguins).[9]

Specific instances. Schemata have slots not only for different attributes and class-inclusion relationships but also for individual instances. For example, the schema for bird may include the information that birds fly and that canaries are a subtype, plus the information that Tweety is a specific instance.

Assessing Schemata

Cohen and Murphy (1984) argue that a schema of this kind can capture all the phenomena family resemblance models can. The retention of instances allows it to reap the benefits of instance models as well. It is also able to handle phenomena that the family resemblance view and instance approach individually have difficulty with, such as conceptual combination (Cohen & Murphy, 1984).[10]

Explicitly representing class-inclusion and constituent relationships in the same structure also allows the schema view to capture the observation that when subjects are asked to list attributes for object concepts, they not only give information about constituent attributes but also almost invariably give information about superordinate (and sometimes subordinate) categories. For example, for piano, many subjects list "a kind of musical instrument" much as they list "made of wood" (Ashcraft, 1978; Komatsu, 1983). Furthermore, by providing a consistent representational structure for both abstracted information and information about instances, the schema view suggests a flexibility between emphasizing particularized or abstracted information, depending on the context of use or acquisition of the concept (cf. Brooks, 1978, 1987; Malt, 1989; Medin, Altom, & Murphy, 1984; Medin & Smith, 1981; E. E. Smith, 1989; Wattenmaker, 1991a).

To the extent that schemata guide the pickup of information about new instances (Neisser, 1975), the schema view should be able to predict effects of expectations about the distributions of instances on concept acquisition (Flannagan et al., 1986). However, making such predictions rests on explicating exactly how schemata guide information pickup and abstract across instances. Although some work on this question has been carried out with regard to the identification of central tendencies (e.g., Medin et al., 1984; Medin, Dewey, & Murphy, 1983; Medin & Smith, 1981; Nakamura, 1985; Nosofsky, Clark, & Shin, 1989; Reed, 1978; Wattenmaker, 1991a, 1991b), results have been tentative and unclear (see also Medin, 1986). Among the variables that may increase the likelihood of extracting central tendencies at acquisition may be the extent to which subjects perceive their task to be a matter of identifying categories (cf. Brooks, 1978) or how tightly structured the categories are and how many instances subjects see (cf. Breen & Schvaneveldt, 1986; Homa et al., 1991; Homa, Rhoads, & Chambliss, 1979; Homa, Sterling, & Trepel, 1981; Omohundro, 1981).

Economy, Informativeness, Coherence, and Naturalness

Because the schema view does not require that all information about all instances be retained, it is no worse off than the exemplar view in terms of economy. In fact, because it predicts that the information picked up (and retained) about new instances is guided by information abstracted from previous instances, the schema view imposes a measure of systematicity on which instances or details are retained. Presumably, this systematicity normally would be recruited in the service of informativeness as well, so that the details retained are those that are nonredundant with the default values that can be inferred.

Coherence, however, poses a problem for the schema view as I have described it here. Once again, it is possible to appeal to the ecological constraint originally mentioned in connection with the family resemblance view. An alternative approach would be to focus on the relationships that schemata encode. A step in that direction was taken by Cohen and Murphy (1984), who point out that schemata may code not only statistical, numerical, or logical relationships between slots and slot values but also functional and causal relationships. Such relationships have the potential for increasing the informativeness of con-

[9] In some versions of the schema view (e.g., Cohen & Murphy, 1984), the relationship of class inclusion itself (and not just inherited slot values) has the status of a default value, that is, the status of being revisable under particular circumstances (but see Minsky, 1975, for a contrary position). The empirical motivation for this qualification comes from Hampton (1982), who demonstrated that at least some subjects consider some category statements to be intransitive. For example, some subjects will agree that chairs are a kind of furniture and car seats are a kind of chair but will not agree that car seats are a kind of furniture. Qualifying the inheritance of slots, slot values, and class inclusion relationships in this manner would make the schema view consistent with the suggestions of Quine (1953) and Wittgenstein (1953; Baker, 1974) on revisability and criteria and distinguishes it from earlier network models (e.g., Collins & Quillian, 1969), in which inheritance of attributes was absolute and supported logical entailment (Cohen & Murphy, 1984).

[10] Connectionist models are sometimes regarded as implementations or instantiations of schema models (Rumelhart, Smolensky, McClelland, & Hinton, 1986, although others regard connectionist models as being most closely related to family resemblance, e.g., Knapp & Anderson, 1984, or exemplar models, e.g., Kruschke, 1992). Connectionist and distributed array models in fact do demonstrate all of the desirable characteristics of schemata, plus a few others (e.g., content addressability). Such models have been the focus of a great deal of attention recently (Estes, 1986a, 1986b, 1988; Estes, Campbell, Hatsopoulos, & Hurwitz, 1989; Gluck, 1991; Gluck & Bower, 1988a, 1988b; Knapp & Anderson, 1984; Kruschke, 1992; McClelland & Rumelhart, 1985; Shanks, 1990, 1991; Smolensky, 1986). However, connectionist and array models do not currently lend themselves to a transparent means for explicating the causal and functional nature of certain relationships that are important for seeing the link between schema and explanation-based views.

cepts as well as explaining what privileges some categories over others.

Although this change from focusing on attributes (or slot and slot values) to focusing on relationships (including functional and causal relationships) among attributes and categories is a relatively simple one, it points the way to a restructuring of the view of concepts. This change in emphasis is what distinguishes the explanation-based views from the similarity-based ones.

Explanation-Based Views

In this section, I review some of the explanation-based views developed by psychologists and other cognitive scientists. However, it is useful to begin with some insights derived from work by philosophers on the nature of reference (i.e., of how words refer to things in the world) and to consider the implications of those insights for psychologists.

The Direct Theory of Reference

Psychologists have traditionally made at least two assumptions about concepts: (a) that they mediate the link between words and the things to which they refer and (b) that they are the representations over which thinking and reasoning take place. These assumptions have had two consequences: (a) Concepts are generally viewed as descriptions of the members of certain categories, and (b) it is assumed that to explicate the nature of concepts is to explicate the nature of meaning. The philosopher Hilary Putnam (1975a, 1975b), however, has argued that (a) concepts do not mediate the link between words and their referents (the things to which they refer) in all cases, (b) although an explication of concepts is part of a theory of meaning, explicating meaning also requires explicating reference, and (c) concepts may be more like theories than (mere) descriptions (see also Rey, 1983, and Schwartz, 1979, for explications of Putnam's views intended for psychologists).

Natural Kind and Nominal Kind Terms

Putnam (1975a, 1975b) argues that people are often able to use words to refer to categories of objects without knowing exactly how those objects are to be described or what it is that holds those objects together in that category. For example, as a child, I read something in which sassafras was mentioned but not described. When I subsequently asked an adult, "What is sassafras?" I successfully referred to something (i.e., sassafras) even though I had no idea what that something was (i.e., effectively had no concept of sassafras). On that occasion, no mental representation mediated between the word *sassafras* and its referent. In such cases, reference is said to be direct (Kripke, 1972). But if reference is direct in cases in which no concept is available, then it is possible that reference is also direct (i.e., unmediated) when a concept is available. The presence of a concept does not guarantee that it mediates the link between a word and its referents. Putnam argues that for a certain class of common nouns, called *natural kind terms,* concepts in fact do not mediate that link.[11]

Natural kind terms label naturally occurring things like water and biological kinds. Two observations are relevant: First, the things labeled by natural kind terms (i.e., natural kinds) are (in our culture) the object of study by scientists or other experts. Second, part of what scientists try to do is discover the true nature of things (e.g., the true nature of the stuff we call *water*) and why those things have the attributes that they do. These two observations led Putnam (1975a, 1975b) to argue that the properties of natural kinds are a matter of empirical (scientific) discovery and not just a matter of linguistic convention (as it is with nominal kind terms such as *bachelor;* see below).

With natural kind terms, there is what Putnam (1975b) calls a "division of linguistic labor" (p. 227): It is left to the experts (i.e., the scientists) to figure out what, for example, water, actually is (i.e., which are the different true instances of water, what it is that those different true instances actually have in common, and why those commonalities emerge). Lay people use natural kind terms like *water* and typically have some beliefs about the kind (i.e., they have a nontrivial concept water), but their concept merely reflects, to a greater or lesser extent depending on the person (and the kind), the experts' beliefs about the kind. This analysis leads to three important implications: (a) Well-developed concepts will be more like theories than simple descriptions, (b) when correct classification is crucial, lay people will defer to the beliefs of scientists, and (c) (lay) concepts of natural kinds do not mediate the link between natural kind terms and their referents. The first implication follows in a straightforward fashion from the claims that experts study and theorize about natural kinds and that those theories eventually spread to other speakers. The second follows from the claim that there is a linguistic division of labor: If some task is the special charge of experts, lay people will defer to the relevant experts in that task. The last, however, bears closer examination.

The possibility of rejecting one's own beliefs about a natural kind in favor of an expert's, according to Putnam, arises because people tend to use natural kind terms referentially (Donnellan, 1966/1977). When one uses a term referentially, he or she means to pick out the stuff so labeled, rather than something meeting a particular description. For example when I use *sassafras* referentially, I mean to refer to actual instances of sassafras, not to whatever meets the description I happen to have in my head of sassafras (which is still pretty minimal and

[11] There is, in fact, some controversy among philosophers over exactly which common nouns are supposed to behave in this fashion. Putnam (1975b) suggests that his ideas apply to most words, including both natural kind terms and artifacts (such as *pencils, chairs*). Although some (e.g., Kornblith, 1980; Losonsky, 1990) have supported Putnam's (1975b) position, others (e.g., Schwartz, 1978, 1980) have argued that the analysis does not apply to artifacts. More significantly, Dupré (1981) has suggested that although Putnam's (1975b) analysis may apply to some natural kind terms, it does not apply to biological kind terms (such as *cat*), at least as such terms are used by biologists. Dupré (1981) argues that the taxonomies biologists construct often involve explicit definitions that violate pretheoretic notions of "same kind of stuff." For the most part, psychologists who have drawn on Putnam's (1975b) analysis seem to have assumed that it does not apply to artifacts but does apply to natural kinds, including biological kinds (cf. Gelman, 1988a, 1988b; Gelman & Markman, 1987; Keil, 1989).

would not uniquely pick out anything). If an expert were to give me an account of sassafras different from one I believe, I would probably change my concept of sassafras, rather than conclude that the expert's word *sassafras* refers to some different kind of stuff than mine. In fact, Putnam (1975a, 1975b) argues that I should be able to accept that scientists may discover that much (or even all) of what I believe to be true of a particular natural kind may be false. But the degree of match between my concept of sassafras and the expert's does not affect the fact that when I use *sassafras* referentially, I refer to the same stuff as does the expert.

Nominal kind terms (Schwartz, 1978, 1979) like *bachelor* or *triangle*, in contrast, tend to be used *attributively* (Donnellan, 1966/1977). Nominal concepts are established by convention rather than empirical discovery (i.e., figuring out what makes something a bachelor does not involve studying various instances of bachelors, except perhaps when first learning the term). When one uses nominal kind terms, he or she generally means to pick out stuff meeting particular descriptions. Thus, when I use *triangle* attributively, I mean to pick out something that meets the description "three-sided planar figure." In contrast to the referential use of natural kind terms (and their codification through discovery), the attributive use of nominal kind terms (and their codification through convention) means that I would find it extraordinarily odd to be told that bachelors in fact are married or that triangles in fact have four sides (see also Kripke, 1979).

Keil (1989) argues that rather than thinking in terms of a strict dichotomy between natural and nominal kind terms, people should recognize a continuum in the degree to which terms tend to be used attributively or referentially, with natural kind terms generally falling toward the referential end and nominal kind terms toward the attributive end (with terms for artifacts such as *pencils* and *cars* perhaps falling somewhere in the middle). I believe that there is also likely to be a continuum in the degree to which people are willing to accept the revisability of attributes and relationships that they believe to be true of the referents of different natural and nominal (and artifact) terms (see Footnote 5).

Implications for a Psychological Theory of Concepts

Although somewhat controversial in some of its particulars (see, for example, Lakoff, 1987b; Salmon, 1981), Putnam's (1975a, 1975b) comments on reference have several implications for a psychological theory of concepts.

First, they suggest that there are different kinds of words (natural kind terms, nominal kind terms, artifact terms), which correspond to different sorts of things. These different sorts of words may be associated with mental representations that encode different sorts of information, are established through different means (discovery vs. convention), and tend to be used in different sorts of ways (referentially vs. attributively). This is a possibility that a number of psychologists have previously explored (e.g., Armstrong et al., 1983; Barton & Komatsu, 1989; Gelman, 1988b; Johnson-Laird, 1987; Keil, 1989; Malt, 1990; Miller & Johnson-Laird, 1976; E. E. Smith, 1988; E. E. Smith & Medin, 1981; but see Gleitman et al., 1983, for the view that there is no systematicity to different sorts of concepts).

Second, Putnam (1975a, 1975b) implies that the information in the mental representation labeled by a natural kind term is not necessarily true of the real-world referents of the term. This possibility is also not completely unfamiliar to psychologists: Neither the family resemblance, the exemplar, nor the schema view requires that the information specified by a concept necessarily be true of the instances of the concept. However, Putnam goes further than any of these positions, allowing for the possibility that the information specified by a concept for a natural kind can be completely wrong (i.e., every attribute is subject to revision). Because this analysis leads to a rejection of the assumption that concepts for natural kinds mediate between words and referents, an assumption that lies at the core of current psychological views of concepts (but see Macnamara, 1991), it is very important for psychologists to determine whether people in fact will accept the revisability of all the information associated with natural kind terms or whether there are certain boundaries to revisability that vary with type of word.

In a closely related point, Putnam argues that the meaning of some words cannot be completely captured by an explication of the concepts that they label. According to Putnam (1975a, 1975b), the meanings of natural kind terms must include not only something like a concept, but something about the extensions of the terms and the relations that hold among the terms, their concepts, and their extensions. This implies, for example, that the meaning of *bird* involves not only the concept bird (and the mental processes that operate on that representation) but also something about the set of birds in the world and how it is that *bird* manages to refer to those birds. As Putnam (1975b) puts it, "'meanings' just ain't in the head" (p. 227). Therefore, the third implication of Putnam's comments is that a study of concepts is only a part of the study of meaning.

Unfortunately, this may pose a problem for psychologists. J. A. Fodor (1980) has argued that cognitive psychology can only study mental representations and the processes that operate on them (a position he calls "methodological solipsism"). Thus, according to J. A. Fodor, issues such as reference, which involves a relationship between mental representations and the outside world, are outside the reach of empirical psychology. Taken together, Putnam's and J. A. Fodor's arguments imply that (a) a theory of concepts cannot explain reference, (b) a theory of meaning (at least of natural kind terms) must provide descriptions of both concepts and extensions and an explanation of reference, (c) psychologists cannot study reference, and therefore (d) psychologists cannot thoroughly study meaning (at least of natural kind terms; Rey, 1983; Rips, 1986; Smith, Medin, & Rips, 1984).

I think that it remains to be seen whether an adequate psychological account for reference (perhaps one that is based on a revised view of concepts) can be developed (see Carroll, 1985; Macnamara, 1991). But these arguments at least point out the need to explore the possibility that reference is not based on the same representations over which thinking and reasoning take place and that factors that are assumed to be relevant for concepts (e.g., the ecological constraint proposed by Rosch, 1978) may actually be relevant to reference instead.

Fourth, Putnam's comments imply that people's mental representations of natural kinds (and perhaps artifacts) are attempts at capturing the nature of those kinds as understood by scientists or other experts. Thus, people's natural kind concepts are not limited to simple lists of attributes, lists of central tendencies on attributes, ranges of values on attributes, or sets of instances. Instead, people's natural kind concepts are (at least when fairly well developed) rather like theories, attempts to explain the distribution of attributes and instances. When those concepts are not well developed, people's concepts (reflecting a division in linguistic labor) may simply include the belief that such an explanation exists and that others in the linguistic community (e.g., scientists or other experts) may know what that explanation is or are working on uncovering it (see also Malt, 1990). With nonnatural kinds, people's beliefs may include the idea that there are no experts and no to-be-discovered explanations, but only linguistic conventions.

Psychologists and other cognitive scientists have developed a number of ideas related to these implications, to which I now turn.

Conceptual Coherence Based on Explanations Rather Than Simplicity

The classical, family resemblance, and exemplar views of concepts assume that concepts cohere (i.e., that certain categories are privileged) because their instances are similar to one another in some fashion (Medin, 1989; Medin & Wattenmaker, 1987; Murphy & Medin, 1985). Thus, these similarity-based views of concepts assume that categorization is a function of a prior judgment of similarity: An item is placed in some category (i.e., is thought to cohere with other members of the category) if it is judged to be similar to either some abstraction or to other members of that category in the relevant fashion. But how is similarity (in the relevant fashion) to be judged? All of the similarity-based views adopt or implicitly assume some variant of Tversky's (1977; Tversky & Gati, 1978) contrast model of similarity (see E. E. Smith, 1989, for an explicit specification of the relationship). Unfortunately, there are certain limitations to the contrast model.

Limitations of the Contrast Model of Similarity

The contrast model argues that similarity is a function of the number of attributes (weighted for relevance or salience) shared and not shared between objects being compared (Tversky, 1977). One limitation of the model is that it assumes the independence of attribute weights and the combination of weights through summing. This allows the contrast model to describe only linearly separable categories. As discussed above, however, this limitation may be addressed by adopting different methods of combining attribute weights.

A second limitation is that the contrast model generally ignores relationships among attributes. A simple enumeration of properties does not properly describe most concepts, which require particular relationships among attributes (e.g., a bird is not a simple collection of bird parts). A straightforward modification would be to treat a relationship between two or more attributes as simply being another attribute (cf. Tversky's, 1977, suggestion that global properties such as symmetry may be among the attributes included). However, recent evidence suggests that attributes and relationships between attributes have different effects on judgments of similarity (Goldstone, Medin, & Gentner, 1991; Medin, Goldstone, & Gentner, 1990). To accommodate these results, the contrast model must at least be modified to accommodate two different kinds of information (i.e., attributes and relations), which have different effects on judged similarity.

A third limitation arises because the contrast model assumes (rather than explicates) the operation of processes that select and weight relevant attributes. But those processes bear the bulk of the explanatory burden. The similarity between two objects depends crucially on which attributes enter into the comparison for matches and mismatches and on the weights assigned to the attributes that are compared. At the extreme, everything is potentially similar to everything else: Attribute comparison sets can be constructed so that any two objects will differ in no attributes and share a potentially infinite number (e.g., existing at this moment in time, capable of being thought about by one person, and another, weighing less than 200 tons, and less than 201 tons; Goodman, 1972; Murphy & Medin, 1985).

One solution to this problem would be to limit the attributes that enter into the comparison for matches and mismatches to just those that are most relevant (Goodman, 1972; see E. E. Smith et al., 1974, for an example). This strategy, however, effectively places the entire burden of determining similarity (and by the assumptions of similarity-based views, categorization) on whatever process assigns relevance weights to instance attributes. Unfortunately, this is not a simple matter of, for example, counting how frequently an attribute occurs among the instances of a concept: The relevance of an attribute varies with stimulus context, task, and the entity in which it adheres (Gati & Tversky, 1984; Ortony, Vondruska, Foss, & Jones, 1985; Tversky, 1977). One particularly interesting example of this problem (particularly given the assumptions of similarity-based views) comes from work by Rips (1989) demonstrating a dissociation between classification and similarity–typicality judgments.

A Dissociation Between Similarity and Categorization

Rips (1989) presented subjects with stories about natural kinds (specifically, animals) and nonnatural kinds (specifically, human artifacts) undergoing certain kinds of changes. Some of these changes Rips (1989) called "essential" (p. 38). In the case of animals, an essential change was one that was a natural part of the animal's development (e.g., early in its life the animal has bird-appropriate characteristics, but as it matures it takes on insectlike characteristics; presumably, such changes stem from something in the animal's genetic structure, which is considered essential to biological natural kinds; cf. Gelman & Wellman, 1991; Keil, 1989). In the case of the artifacts, an essential change was a change in intended function. (Barton & Komatsu, 1989; Losonsky, 1990; Richards, 1988; Richards & Goldfarb, 1986; Richards, Goldfarb, Richards, & Hassen, 1989, also argue that intended functions are essential to artifacts; but see Malt & Johnson, 1992.)

Other changes Rips (1989) called "accidental" (p. 39), where *accidental* referred not to the deliberateness of the change, but to the fact that the change did not involve the essential nature of the object undergoing the change. For example, in one story, an animal that has certain physical and behavioral characteristics typically associated with birds (but is not specifically labeled as being a bird) eats vegetation contaminated by hazardous wastes. After a while, it undergoes certain accidental changes: The bird-appropriate physical and behavioral characteristics change to forms more closely associated with insects (although the story also specifies that the changed animal produces normal offspring, suggesting an intact "essence"). Accidental changes in an artifact were changes in physical characteristics (e.g., color, parts) that did not affect the artifact's function.

Rips (1989) found that for both natural kinds and artifacts, accidental changes affected classification judgments less than similarity and typicality judgments, whereas essential changes affected classification judgments more than similarity and typicality judgments. For example, the bird that had undergone an accidental change was likely to be categorized as a bird but was perceived to be quite dissimilar to (other) birds and not typical of a bird. On the other hand, in the essential-change condition, the animal (during its birdlike stage) was considered very similar to birds and very typical of a bird but was not considered likely to be a bird.

These experiments suggest that judgments of categorization may differ systematically from judgments of similarity or typicality. This implies that the attributes or attribute weights that are used to determine similarity or typicality are not the same ones used to determine categorization. Therefore, it is unlikely that categorization is based on prior judgments of similarity–typicality, or exactly the same factors as judgments of similarity–typicality. Our view of concepts needs to provide for the two kinds of judgments to proceed independently. In particular, certain attributes (i.e., accidental ones) play a greater role in judgments of similarity or typicality than in judgments of categorization, but other attributes (i.e., essential ones) play a greater role in categorization judgments than in similarity or typicality judgments. But the attributes that are accidental and essential are different for different kinds of concepts. Although concepts appear to be coherent, our use of concepts has a great deal of flexibility.

Characterizing the Explanation-Based View

The explanation-, or knowledge-, based view of concepts tries to explain the simultaneous properties of coherence and flexibility by arguing that the specification of a concept includes information about how that concept is related to other concepts (or how its instances relate to other objects) and about the relationships—especially the functional, causal, or explanatory relationships—that hold among the attributes associated with its instances. For example, the concept piano may include the information that people typically sit on a bench to play it (i.e., a relationship to other concepts). The concept bird may include the information that birds have a certain genetic structure that under normal conditions expresses itself in having wings, feathers, and so on (i.e., relationships among attributes associated with its instances).

Perhaps the clearest description of the modal form of the explanation-based view is found in a quote from Keil (1989):

> [In this view] concepts are construed as intrinsically relational sorts of things. They are not isolated entities connected only in the service of propositions. No individual concept can be understood without some understanding of how it relates to other concepts. Concepts are not mere probablistic distributions of features or properties, or passive reflections of feature frequencies and correlations in the world; nor are they simple lists of necessary and sufficient features. They are mostly about things in the world, however, and bear nonarbitrary relations to feature frequencies and correlations, as well as providing explanations of those frequencies and correlations. If it is the nature of concepts to provide such explanations, they can be considered to embody systematic sets of beliefs—beliefs that may be largely causal in nature. (p. 1)

The explanation-based view is not inconsistent with the schema view (Cohen & Murphy, 1984). In fact, I believe that the explanation-based view is best seen as an elaboration of the schema view. The difference between the schema view and the explanation-based view is one of focus, with the explanation-based view having an increased emphasis on the inclusion of (in particular, causal and explanatory) relationships among schemata and slots (i.e., among different concepts and the different attributes of a given concept).

Empirical Support for the Explanation-Based View

The explanation-based view suggests that our intuitions of conceptual coherence and our flexibility in using concepts stem from an understanding of the explanatory relationships present. Details of just how that information is used to achieve coherence and flexibility have not been thoroughly worked out, and specific predictions made by the explanation-based view are difficult to identify. However, some general comments can be made.

For example, because the explanation-based view (like the schema view) adds to the kinds of information that the family resemblance and exemplar views include in conceptual representations, it has no problems with the phenomena for which those views provide an account. To argue for the explanation-based view over the family resemblance and exemplar views (or the schema view) therefore requires evidence that supports the crucial role of explanations in conceptual phenomena. Recently, some relevant evidence has been found.

Medin, Wattenmaker, and Hampson (1987) reported several experiments that involved cartoon creatures specifically constructed to fall into different family-resemblance-based categories. Subjects in these experiments sorted pictures of these creatures into categories. A variety of stimulus and instructional manipulations failed to induce subjects to sort on the basis of family resemblance. Only when the interproperty relations were made salient to the subjects (i.e., when subjects were made aware of the underlying principles responsible for, or that explained, the family resemblances) did such classification take place. Similarly, Wattenmaker et al. (1986) and (to a limited extent) Nakamura (1985) found that the ease with which people are able to learn different categories is affected more by the activation of possible underlying explanations of the interproperty relationships than by linear separability, and Pazzani (1991) found that the relative difficulty of learning conjunctive versus disjunctive concepts is affected by their consistency with subjects' prior causal knowledge.

Also supporting the important role of explanations in using conceptual knowledge, Murphy (1988, 1990) and Medin and Shoben (1988) found that subjects' understanding of adjective–noun and noun–noun combinations could be fully accounted for only if subjects' world knowledge (e.g., knowledge about how different kinds of objects interact with one another) were allowed to influence attribute selection and weighting. Kunda, Miller, and Claire (1990) reported similar results when subjects were asked to judge combinations of social concepts (e.g., Harvard-educated carpenter).

Economy, Informativeness, Coherence, and Naturalness

According to the explanation-based view, instances of a concept cohere because some set of relations links them with one another, and not with noninstances. Similarly, the set of instances of a concept is regarded as forming a natural category because of the availability of an explanation that holds them together. For example, living objects with wings, hollow bones, and feathers form a natural category because all of those attributes may be (by theory) manifestations of a single underlying genetic structure; (by theory) sharing that underlying genetic structure is what gives the category coherence.

The availability of information about how instances of the concept relate to the world and of the relations (particularly explanatory relations) that hold among attributes of the instances also helps increase the informativeness of explanation-based representations: Such relations help one decide exactly how he or she should go beyond the information given with a particular new instance or with an old instance in a new situation. For example, if a person was told only that some newly discovered mammal lived in the ocean, he or she would probably predict its limbs to be flipperlike rather than leglike (because of his or her understanding of the requirements of locomotion in that environment), although leglike limbs are much more common among mammals generally. Similarly, relational information is taken to explain how it is that attribute relevance varies with context or with different tasks (although, again, the specific means by which it does so are still somewhat unclear).

It may be that these strengths, however, are gained at the expense of economy. The explanation-based view argues that a well-developed concept includes information about the relations among instances of the concept and of other things in the world. If totally unconstrained, this could mean that at the extreme, through explanatory links to other concepts, a single concept might incorporate almost all information available to the person, an issue linked to the knowledge access or knowledge representation and "frame" problems in artificial intelligence (Haugeland, 1985). More generally, in the explanation-based view, the burden of understanding coherence, naturalness, informativeness, and economy has been shifted from a description of how concepts are constrained to a description of how explanations are constrained (Medin & Wattenmaker, 1987; Murphy & Medin, 1985; Wattenmaker et al., 1988). Progress in describing these constraints has been somewhat limited, but several different approaches have been explored.

Variants of the Explanation-Based View

Variants of the explanation-based view have been proposed by researchers investigating metaphors (Lakoff, 1987a, 1987b), reasoning (e.g., Johnson-Laird, 1983), concept development (e.g., Carey, 1985; Gelman, 1988a, 1988b; Gelman & Markman, 1987; Keil, 1986, 1987, 1989), and knowledge representation in computers (e.g., Michalski, 1989; Mitchell, Keller, & Kedar-Cabelli, 1986; Schank, 1982; Schank, Collins, & Hunter, 1986) as well as by those studying adult human concepts (e.g., Medin, 1989; Medin & Ortony, 1989; Medin & Ross, 1989; Medin & Wattenmaker, 1987; Murphy & Medin, 1985; Rips, 1989). I review four variants below.

Explanation-Based Approaches: Psychological Essentialism

Medin and Ortony (1989; see also Keil, 1989; Malt, 1990; Wattenmaker et al., 1988) argue for an idea that they call *psychological essentialism.* Medin and Ortony distinguish psychological essentialism from metaphysical essentialism, the problematic claim that things are what they are by virtue of possessing some internal essence (e.g., that rocks are rocks by virtue of possessing "rock essence").[12] Psychological essentialism does not claim that things have essences, but rather that people's representations of things may reflect a belief that things have essences (e.g., that the representation rock includes a belief in a rock essence, although one may not know precisely what that essence is).

If a person's representation of an object includes the belief that the object is an object of a particular kind by virtue of possessing an essence for that kind, it would explain why people often seem to believe in necessary and sufficient definitions (e.g., McNamara & Sternberg, 1983). However, Medin and Ortony (1989) are clear that they do not intend psychological essentialism to merely recapitulate the classical view, or the classical-plus-family-resemblance dual-representation approach. In fact, Medin and Ortony's proposal for psychological essentialism consists of three distinct claims about the information that constitutes a concept: first, that such information lies on a continuum of accessibility, second, that human perceptual and conceptual systems have evolved to maximize success in identifying the information that is less accessible, and, third, that in some cases the information that constitutes a concept consists in part of an "essence placeholder" (p. 189).

Continuum of accessibility. Medin and Ortony (1989; Keil, 1989) suggest that the attributes associated with the instances of a category range on a continuum from deep, relatively inaccessible attributes to more accessible surface ones. More accessible attributes tend to be perceptual in nature (e.g., having a particular color or certain parts), whereas less accessible attributes are more abstract (e.g., having a particular molecular structure or function) or are more in the nature of "background" knowledge (e.g., how people interact with instances of the concept). Medin and Ortony predict that unless specifically

[12] To see the problem, note that the same object may be described at different times as being a rock, a paperweight, or an ashtray. But each description suggests a different essence. So which is the true essence of that object? The possibility that different essences may be possessed at different times by the same object undermines the plausibility of possession of internal essences as an explanation for what makes a thing what it is.

instructed to the contrary, the accessible, easily perceptible attributes are acquired first, with the less accessible, deeper attributes acquired only as a fuller understanding of the concept develops (this may also hold true developmentally; see below). More important, unlike in the approach combining family resemblance and classical views, Medin and Ortony (and Keil, 1989) argue that surface and deep attributes are inextricably linked: The easily perceptible attributes of the instances of a category are often constrained by, and in some cases generated by, the deeper, more central attributes. For example, a deep attribute of whales is their genetic structure. But this relatively inaccessible attribute in turn is causally responsible for such easily perceptible attributes as their breathing air, having a tail and flippers.[13]

Evolution of perceptual and conceptual systems. Medin and Ortony (1989) suggest that the perceptual and conceptual systems of organisms have evolved in such a way as to be sensitive to just those easily accessible attributes that are likely to lead them to grasp the deeper, underlying attributes. This claim has two implications. First, it implies that the attributes that are accessible to our perceptual system are generally precisely those that are linked to the deeper attributes. This explains why it is often useful to categorize on the basis of perceptual similarity: Given the evolution of our perceptual systems, such categorizations will usually not go far wrong; appearances do not usually deceive (although there are exceptions, such as whales). It also explains why the ecological constraint proposed by Rosch and Mervis (1975; Neisser, 1987; Rosch, 1978) gets as far as it does.

Second, this claim about the evolution of our conceptual system implies that our normal approach to learning new categories should be one that not only facilitates the extraction of underlying attributes, but that allows us to recover from and revise initial classifications that are based on easily accessible attributes. Two observations are relevant. First, as Medin and Ortony (1989) note, L. B. Smith (1989) has found that young children tend to classify on the basis of global (perceptual) similarity (but see Ward, Vela, & Hass, 1990). Such classification, based as it is on superficial attributes, may not give the best basis for classification (which presumably is provided by deeper, explanatory attributes), but it is also unlikely to go completely astray, as classification on a single, wrong, superficial attribute might.[14] Second, as noted in connection with the exemplar view, there is good evidence that information about specific instances of a concept is often retained. This makes sense: If easily accessible attributes do not give the best basis for classification but provide clues for a deeper, explanation-based categorization, it is prudent to retain as much information as possible initially rather than to prematurely discard all but summary information (see also Wattenmaker, 1991a). Such summary information may later turn out to be irrelevant, or worse, misleading (Medin & Ross, 1989). On the other hand, the possibility of eventually extracting deeper, explanatory relations implies that abstracted information is a crucial aspect of well-developed conceptual representations.

Essence placeholder. Finally, Medin and Ortony (1989) apparently mean to identify the deepest, most central attributes of a concept (at least some cases) as an essence placeholder. In some cases, the deeper attributes may be a simple list of attributes singularly necessary and collectively sufficient for being an instance of the concept, much as claimed by the classical view. In contrast to the classical view, however, Medin and Ortony argue that in other cases the necessary and sufficient attributes may be a consequence of, rather than constituent of, the essence placeholder. Medin and Ortony also argue that the essence placeholder may be a set of inchoate beliefs that are more like a theory of what it is to be an instance rather than a simple list. Characterized in this fashion, the relationship between Medin and Ortony's proposal and Putnam's, (1975b), description of the mental representation of natural kinds becomes apparent. (Medin and Ortony also adopt Putnam's, 1975b, notions of a linguistic division of labor and suggest that in still other cases, the essence placeholder may not include a theory, inchoate or not, of what it is to be an instance of a category but may include, rather, a belief that other individuals either know or pursue the development of a such a theory.)

By allowing the essence placeholder to take these different forms, Medin and Ortony (1989) are able to provide a unified description for some of the different kinds of concepts identified above. For example, allowing for the essence placeholder to be a list of individually necessary and collectively sufficient attributes enables the explanation-based view to capture nominal kind concepts such as triangle or septuagenarian that are acknowledged to be definitions in a conventional sense and

[13] Psychological essentialism clearly descends from the classical-plus-family-resemblance dual-representation approach sometimes called the *core-plus-identification function model* (see discussions in Rey, 1983, 1985; E. E. Smith, Medin, & Rips, 1984). Another descendent is E. E. Smith's (1989) core-plus-prototype approach. E. E. Smith argues that concepts have two of the classical-plus-family-resemblance dual-representation approach components, one relatively accessible and used for rapid identification (the prototype) and the other less accessible but the source of conceptual stability and the "ultimate arbiter" (p. 510) of categorization (the core). Like Medin and Ortony (1989), E. E. Smith does not mean for the less accessible aspect of the concept (at least with nonnominal concepts) to be a definition. Unlike Medin and Ortony, however, E. E. Smith makes no commitment regarding a causal relationship between the less and more accessible aspects of a concept. Furthermore, E. E. Smith suggests thinking of the core and prototype as being represented by schemata and makes no commitment regarding a theorylike structure to the core. The most innovative aspect of E. E. Smith's proposal, however, does not lie in the kind of conceptual representation he describes but in the way in which he combines the core-plus-prototype representation with a general model of induction to explain a variety of results, including Rips's (1989) demonstration of a dissociation between judgments of similarity–typicality and categorization.

[14] On the other hand, adults (at least in Western cultures) seem to be disinclined to use this approach, typically preferring to adopt an analytic approach, and to classify on the basis of one or at most a few easily accessible attributes when faced with a category-learning task (Brooks, 1978; Kemler Nelson, 1984; Medin, Wattenmaker, & Hampson, 1987; Ward & Scott, 1987). This result, however, may reflect subjects' expectations about the kinds of categories experimenters are likely to construct rather than the strategies they use with novel, natural categories. In particular, subjects may not expect categories in psychological experiment to have a deep structure (e.g., underlying relations among attributes) and so may concentrate only on individual, easily perceptible attributes.

whose labels tend to be used attributively. Allowing for the essence placeholder to be more like an inchoate theory, perhaps of genetics or molecular structure, or a belief that other speakers of the language have or are developing a theory enables the explanation-based view to capture natural kind concepts (cf. Malt, 1990). Allowing for the essence placeholder to specify a function, and something about the relationship between that function and certain physical attributes, enables this view to capture concepts for artifacts such as chair or tire (Barton & Komatsu, 1989).

Conceptual coherence and constraints on attributes. The essence placeholder is what provides coherence to a category: It explains why the instances go together in one way and not another. The essence placeholder also provides the theoretical framework and background knowledge to constrain and determine the relevance of easily accessible attributes. In some cases, the constraint imposed on the less accessible attributes is relatively direct and strong. For example, a particular genetic structure (the essence placeholder) directly and strongly constrains certain physical attributes, like having hair (a superficial attribute). In other cases, the constraint is less direct or less strong. For example, information about how people typically interact with instances of the concept chair (i.e., people sit on them) helps determine the range of possible materials out of which chairs may be made (i.e., it must be something capable of supporting a certain amount of weight).

The manner in which the essence placeholder determines the effects of context on the relevance of attributes, however, is much less clear. One example given by Murphy and Medin (1985) concerns the relevance of the attribute flammable to the concept currency. Typically, flammable is an attribute of low relevance to currency. However, background knowledge that people try to save valuables that burn when their house is on fire leads to flammable being highly relevant to currency in such a context.

Note that the currency example suggests that the essence placeholder for a concept either includes general knowledge not directly concerned with the particular concept (i.e., encyclopedic or practical knowledge such as the desire to save valuables in a fire, as opposed to the dictionary knowledge specific to a concept; see Miller, 1978) or is the point of interaction between the concept and general knowledge. In either case, Medin and Ortony (1989) do not specifically address the problem of how the essence placeholder, provides explanations that unify and constrain concepts or is itself to be constrained.

Explanation-Based Approaches: Idealized Cognitive Models

The position that concepts are distinct from but interact with general knowledge is made somewhat clearer by Lakoff (1987a, 1987b) in his suggestion that concepts are best understood as idealized cognitive models (ICMs). Lakoff emphasizes that ICMs are idealized: built on background circumstances or assumptions that may not hold in the real world. Otherwise, at least some of Lakoff's ICMs are much like Medin and Ortony's (1989) essence placeholders: They either may specify necessary and sufficient criteria for inclusion in a category or may be a kind of theory about the instances of the category.

Interaction of conceptual and background knowledge. One example of the former that Lakoff uses is that of bachelor, a well-defined concept that nonetheless can give rise to typicality effects. Lakoff proposes that our model of a bachelor is that of an unmarried adult man. But this model is idealized and makes certain background assumptions about marriage and what constitutes marriageability that do not always obtain in the real world. For example, those assumptions do not take into account the existence of priests, homosexuality, or long-term, romantic, cohabiting relationships. According to Lakoff, it is this lack of fit between background assumptions and the real world that occasionally arise when we use the concept, rather than any fuzziness in bachelor itself, that is the source of typicality effects.

Thus Lakoff's ICM approach emphasizes, as do all explanation-based approaches, that not all the effects to be observed in the use of concepts can be explained by considering concepts in isolation. Such effects often emerge, Lakoff (1987a, 1987b) claims, through an interaction between the information specific to a particular ICM and background knowledge.

Lakoff (1987a, 1987b) also suggests that typicality effects do sometimes occur as a direct result of the structure of the particular concept. Such is the case with so-called *radial* categories, which cannot be captured by a single cognitive model but instead appear to involve a cluster of related models. One example is the concept mother. The multiplicity of our understanding of mother is conveyed by the following sentences (Lakoff, 1987b):

> I was adopted and I don't know who my real mother is.
>
> I am not a nurturant person, so I don't think I could ever be a real mother to any child.
>
> My real mother died when I was an embryo, and I was frozen and later implanted in the womb of a woman who gave birth to me.
>
> I had a genetic mother who contributed the egg that was planted in the womb of my real mother, who gave birth to me and raised me. (p. 75)

Conceptual coherence and constraints on attributes. Lakoff (1987a, 1987b) describes a variety of ICMs that have different sorts of structure, and the specific constraints and accounts for coherence depend on the particular kind of ICM being discussed. Some (such as the model for bachelor) have the shape proposed by the classical view and are constrained and given coherence by having a necessary and sufficient set of attributes (although why the particular set of attributes is selected in the first place is not addressed). Those with a nonclassical shape, such as mother, depend on something like a theory or explanation or general principles of extension from central cases. In the latter case, Lakoff acknowledges that the theory or general principles do not predict the range of the category, although they motivate membership in the category. In many ways, this is a fair assessment of explanation-based approaches in general.

Explanation-Based Approaches: Mental Models

A third explanation-based variant is Johnson-Laird's (1980, 1981, 1983, 1987) mental models approach. Johnson-Laird (1983) suggests that representations corresponding roughly to concepts may be understood most profitably as schemata in long-term memory. Emphasizing an aspect of schematic knowledge mentioned by Rumelhart (1980) but not focused on by the psychological essentialism or ICM approaches, Johnson-Laird argues that such schemata include information about how concepts are to be used to construct (in working memory) mental models. These mental models encode how instances of the concept interact with one another and with other objects and forces in the real world.[15]

Constructing models in working memory. Johnson-Laird (1980, 1981, 1983, 1987) argues that conceptual effects (e.g., typicality effects, judgments about attribute relevance) do not derive directly from the conceptual representations stored in long-term memory but rather from the mental models constructed in working memory.[16] Because implicit inferences that are based on background knowledge and (perhaps inchoate) theories about how the world works affect the construction of the mental models, context effects emerge in a natural (if still somewhat underspecified) fashion from a mental model approach. An example given by Johnson-Laird (1983, p. 234) for the concept fly is particularly clear. Therefore, I use that concept, although it is for an action rather than an object and the analysis centers around understanding sentences rather than simple concepts.

Consider the sentence:

> Alcock and Brown were the first to fly *X* from the USA to Ireland.

What sort of thing, Johnson-Laird (1983) asks, can *X* be, and how can one understand the relationship between *X* and fly? At first, one may be inclined to restrict *X* to a vehicle of some sort and claim that *X* is what does the flying. But note that depending on what kind of vehicle *X* is, the relationship between *X* and fly changes. Contrast

> Alcock and Brown were the first to fly an airplane from the USA to Ireland

with

> Alcock and Brown were the first to fly a bicycle from the USA to Ireland.

To make things more complicated, consider

> Alcock and Brown were the first to fly the Atlantic from the USA to Ireland.

One could say simply that the concept fly enters into three different kinds of relationships. In that case, given the structure x fly y, one would have three possibilities:

(1) x controls y's path through the air
(2) x takes y in an aircraft
(3) x travels in the air over y (Johnson-Laird, 1983, p. 234)

But if one settles on that, this sentence

> I saw the Azores flying the Atlantic (Johnson-Laird, 1983, p. 235)

is ambiguous: It can be read either as saying that I saw the Azores as I was flying over the Atlantic or that I saw the Azores as they were flying over the Atlantic (Johnson-Laird, 1983, p. 235). People rarely recognize that ambiguity, however.

This could be explained if it is proposed that in understanding the combination of concepts presented in the sentence, people construct a mental model of the situation being described. Their construction of the mental model is guided by their knowledge about the concepts referred to by the words and by their knowledge about what sorts of things the objects referred to in the sentence can do. Johnson-Laird (1983) argues that it would be hopeless to try and specify ahead of time what sorts of things can and cannot fly. Instead, the ambiguity is resolved by an implicit inference (which comes about in the course of the construction of the mental model) about what sorts of things islands are (i.e., large land masses rooted at the bottom of the ocean) and what that implies about their ability to fly.

Conceptual coherence and constraints on attributes. By focusing on the construction of representations in working memory, the mental models approach gives inferences an explicit role in conceptual effects. Such inferences allow the mental models approach to account for a wide range of data, including coherence (but see Ford, 1985; Oden, 1987; Rips, 1986, for criticisms). These inferences build on background knowledge, explanations, and theories about how instances of a category interact with one another and with other objects and forces in the real world. Such knowledge, Johnson-Laird (1980, 1981, 1983, 1987) argues, is part of the schemata stored in long-term memory, that is, the concepts. Thus, no distinction between information particular to specific concepts and general knowledge is maintained in the mental models approach, leaving the range and kinds of information that a concept can include unconstrained.

Explanation-Based Approaches: The Two-Tier Approach

In his two-tier approach, Michalski (1989) is more explicit than most of the others in arguing that understanding of a concept in a particular context is the result of an interaction between two components. In Michalski's approach, these two components are the base concept representation (BCR) and the inferential concept interpretation (ICI). The BCR is a representation in long-term memory that includes both specific and general information about the concept. The specific information Michalski suggests includes examples, counterexamples, and exceptions. The general information about the concept includes "typical, easily definable, and possibly context-independent assertions about the concept [that] tend to capture the principle, the ideal or intention behind a given concept" (Michalski, 1989, p. 123). It may either be derived from the application of inferential processes on the specific information or be obtained through direct input (e.g., someone saying, "tires are typically made of rubber"). The BCRs, however, seem to more limited in informational content than the long-term memory

---

[15] In this way, Johnson-Laird tries to capture, albeit with a mental representation, some of Putnam's (1975b) comments about extensions and reference being part of the meaning of a word while trying to avoid the problem posed by J. A. Fodor's (1980) argument regarding methodological solipsism (cf. Johnson-Laird, 1983).

[16] A similar position is taken by Barsalou (1983, 1985, 1989). Barsalou argues that concepts should be understood as constructs in working memory rather than as representations in long-term memory, although he does not emphasize the extensional modeling issue, as Johnson-Laird does.

schemata proposed by Johnson-Laird (1983), in that they do not carry extensive information about how instances interact with one another and other objects or forces. Such information instead is embodied by the ICI.

Michalski (1989) describes the ICI as a process that applies inferential methods to the BCR, making use of relevant background knowledge to yield an interpretation of the concept appropriate to the context. Unlike the inferential mechanisms described or implied by other supporters of the explanation-based view, the ICI is not a general inferential procedure. Instead, Michalski claims that the ICI is specific to each concept, in that the ICI incorporates "metaknowledge about the concept, that is, which properties of the concept are crucial and which are not in a given context, what transformations are allowed on the BCR, and how these properties or transformations can vary among instances of the concept" (p. 129). Thus, although Michalski may apportion and locate certain kinds of conceptual information differently than others proposing explanation-based approaches, the kind of metaknowledge he describes as necessary for understanding concepts does not differ from the other explanation-based approaches described.

Michalski (1989) does not explicitly address the issue of conceptual coherence. Because the kind of metaknowledge incorporated into Michalski's ICI is thought to be the source for coherence in other explanation-based accounts, it seems reasonable to assume that the ICI plays such a role in the two-tier account.

By limiting the BCR to "typical, easily definable, and possibly context-independent assertions," (p. 123) Michalski (1989) places a weak constraint on the attributes that constitute the BCR. But beyond specifying that the ICI contains, or has access to, information that is specific to each concept, Michalski does not seem to place any constraint on the kind of information that the ICI contains or has access to. Insofar as Michalski requires both BCR and ICI for each concept, Michalski's approach, like the mental models approach, therefore effectively places no real constraint on concepts.

Unaddressed Issues in the Explanation-Based View

In the similarity-based views of concepts, the judgment to include an item in a particular category (i.e., categorization) was assumed to be a straightforward consequence of the similarity between the to-be-judged item and the members of the category. The necessary similarity judgment was assumed to be a straightforward consequence of the degree of match and mismatch between the to-be-judged item and a stable mental representation of the category. Concepts were assumed to be these stable mental representations. But the flexibility with which one is able to use concepts (coupled with the coherence of conceptual categories) has made it reasonably clear now that judgments of categorization and similarity are not as simple as had been assumed. It is even somewhat unclear whether concepts are stable representations.

Concepts as Temporary Constructions

Some versions of the explanation-based view (e.g., Barsalou, 1987; Michalski, 1989; in some ways, Johnson-Laird, 1983) may be interpreted as suggesting that what one usually thinks of as concepts (i.e., the representations responsible for typicality effects, etc.) are representations constructed (in working memory) in a particular context. Thus concepts are not stable representations but rather are emergent from the application of certain operations on a base of information in long-term memory.

Not very much is currently known about these operations, but because they must be sensitive to certain global as well as specific contextual factors, they cannot simply retrieve certain prespecified information and must instead take the form of inference rules. This means that concepts only exist after certain kinds of inferences have taken place. A full explication of concepts therefore requires an explication of these inference rules, the kinds of information they operate on, the kinds of information they yield, and the contexts that trigger their operation. But this approach violates the assumption that conceptual representations can be discussed independently of the processes that operate on them.

Constraining Theories, Distinguishing General and Conceptual Knowledge

Other versions of the explanation-based view (e.g. Keil, 1989; Medin & Ortony, 1989, some of Lakoff's, 1987b, ICMs), following the tradition of separating representation and process, diminish the contribution of specific inferential processes and account for conceptual phenomena by proposing (or assuming) the operation of relatively simple procedures on stable but enriched conceptual representations (see E. E. Smith, 1989, for a rather detailed description of this strategy). These accounts suggest that mature concepts typically include theories or explanations about the relationships among different concepts (or their instances) or different attributes of the instances of a given concept. But this raises two problems.

The first problem is that constraints must then be specified on these theories or explanations. Without such constraints, one could come up with explanations that motivate practically any category. Some progress along these lines has been made, however. For example, Keil (1989) suggests that the naive theories that constitute the psychological essence should seek to explicate causal relations, to maximize the number of elements of a concept that are causally linked to one another, and to cluster such causal relations hierarchically (because that kind of building block structure gives economy of representation and stability of substructures; Simon, 1981). However, more work in this area is needed before a clear distinction can be made between the kinds of explanations that can and cannot serve as the psychological essence of a concept. More work is also needed before these theories can predict the categories that will be considered coherent, rather than simply motivate the coherence of known categories (Goodman, 1972; Osherson, 1978).

The second problem is that such theories or explanations can be very difficult to distinguish from background or general knowledge. For example, is information about how people typically interact with chairs and the implications of those actions for the materials chairs can be made of information that is specifically part of the concept chair or part of a general pool of background knowledge? If general or background knowledge and the theories and explanations that are part of a concept in this view cannot be distinguished, then there is no constraint on the amounts or sorts of information that are part of a single concept (Medin & Wattenmaker, 1987; Murphy & Medin, 1985). This leads to the possibility that any piece of knowledge a person has may play a role in his or her use of certain concepts. That kind of a global accessing capacity would make even simple effects for such concepts, in J. A. Fodor's (1983), words, isotropic. As Murphy (1988) points out, one of J. A. Fodor's (1983) main theses about isotropic processes is that they are difficult, perhaps in practice impossible, to fully understand. This suggests that making a distinction between information that can or cannot, or between tasks that do or do not, access general knowledge is important to future progress on understanding concepts.

Conclusion

Through the 1970s, psychologists' views of categorization were based on similarity. Certain objects were thought to be categorized together because they were sufficiently similar, either to one another or to some abstract specification of a definition or central tendency. The concepts on which these categorizations were based were therefore viewed as mental representations of definitions, central tendencies, instances, or some combination of the latter two. With the 1980s, however, came a growing awareness of the inadequacy of such similarity-based views and of philosophical work that suggested that the meanings of certain terms include a specification of their real-world extensions and that the mental representations associated with such terms are much more like theories than simple lists.

This led several cognitive scientists to develop a new view of concepts, emphasizing the role of explanations and inferences. In this explanation-based view, concepts are thought to include information about their relations to one another and about the relations (particularly, causal relations) among the attributes displayed by their instances, in addition to the abstracted, central tendency information proposed by the family resemblance view and the instance-specific information proposed by the exemplar view.

Although there are several variants of the explanation-based view that have yet to be integrated, this new view is very promising. It gives interesting insights into the problems of context effects, conceptual combination, conceptual coherence, and conceptual naturalness and into the relationship between "superficial" (perceptually based), or immature, concepts and deeper, "cognitive," or more mature, ones. But it also raises a number of difficult questions: Can the traditional distinction between dictionary and encyclopedic, or general, knowledge be maintained? If the distinction is maintained, how do concepts and general knowledge interact? What are the different kinds of inferential processes that operate on concepts, how do they do so, and can information derived through inferences be distinguished from information represented by concepts? What is an explanation, and exactly how does explanatory information affect our judgments about similarity, typicality, and relevance in different contexts and about the combination, coherence, and naturalness of concepts? If categorization is not simply a matter of similarity or typicality, what is the relationship between them? Can a purely psychological theory account for the full range of referential phenomena, and if so, what would such a theory look like?

The explanation-based view casts doubt on the assumption that concepts can be understood independently of the inferences they enter into. It may even cast doubt on the assumption that concepts are stable representations. Abandoning these two assumptions makes the study of concepts enormously more complicated, for it suggests that what concepts are can be understood only in the context of a general and complete theory of cognition. This means that one of the first issues future research on concepts will need to clarify is whether the study of concepts is a coherent and relatively independent component of the study of cognition or whether the study of concepts can only take place as a well-enmeshed piece of a larger study of cognition in general.

References

Anderson, J. R. (1978). Arguments concerning representations for mental images. *Psychological Review, 85,* 249–277.

Anderson, J. R. (1980). Concepts, propositions, and schemata: What are the cognitive units? In *Nebraska symposium of motivation* (Vol. 28, pp. 121–162). Lincoln: University of Nebraska Press.

Anderson, J. R. (1990). *The adaptive character of thought.* Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1991a). The adaptive nature of human categorization. *Psychological Review, 98,* 409–429.

Anderson, J. R. (1991b). Is human cognition adaptive? *Behavioral and Brain Sciences, 14,* 471–517.

Anderson, J. R., Kline, P. J., & Beasley, C. M., Jr. (1979). A general learning theory and its application to schema abstraction. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 13, pp. 277–318). San Diego, CA: Academic Press.

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

Ashby, F. G., & Lee, W. W. (1991). Predicting similarity and categorization from identification. *Journal of Experimental Psychology: General, 120,* 150–172.

Ashcraft, M. H. (1978). Property norms for typical and atypical items from 17 categories. *Memory & Cognition, 6,* 227–232.

Baker, G. P. (1974). Criteria: A new foundation for semantics. *Ratio, 16,* 156–189.

Barresi, J., Robbins, D., & Shain, K. (1975). Role of distinctive features in the abstraction of related concepts. *Journal of Experimental Psychology: Human Learning and Memory, 1,* 360–368.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & 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* (pp. 101–140). New York: Cambridge University Press.

Barsalou, L. W. (1989). Intraconcept similarity and its implications for interconcept similarity. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 76–121). New York: Cambridge University Press.

Barsalou, L. W., & Sewell, D. R. (1985). Contrasting the representation of scripts and categories. *Journal of Memory and Language, 24,* 646–665.

Bartlett, F. C. (1932). *Remembering.* London: Cambridge University Press.

Barton, M. E., & Komatsu, L. K. (1989). Defining features of natural kinds and artifacts. *Journal of Psycholinguistic Research, 18,* 433–447.

Bellezza, F. S. (1984a). Reliability of retrieval from semantic memory: Common categories. *Bulletin of the Psychonomic Society, 22,* 324–326.

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

Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review, 88,* 354–364.

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution.* Berkeley: University of California Press.

Bourne, L. E., Jr. (1970). Knowing and using concepts. *Psychological Review, 77,* 546–556.

Bourne, L. E., Jr. (1982). Typicality effects in logically defined categories. *Memory & Cognition, 10,* 3–9.

Bourne, L. E., Jr., Dominowski, R. L., & Loftus, E. F. (1979). *Cognitive processes.* Englewood Cliffs, NJ: Prentice-Hall.

Breen, T. J., & Schvaneveldt, R. W. (1986). Classification of empirically derived prototypes as a function of category experience. *Memory & Cognition, 14,* 313–320.

Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169–211). Hillsdale, NJ: Erlbaum.

Brooks, L. R. (1987). Decentralized control of categorization: The role of prior processing episodes. In U. Neisser (Ed.), *Concepts and conceptual development* (pp. 141–174). New York: Cambridge University Press.

Brown, R. (1958). How shall a thing be called? *Psychological Review, 65,* 14–21.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking.* New York: Wiley.

Busemeyer, J. R., Dewey, G. I., & Medin, D. L. (1984). Evaluation of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 638–648.

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

Cantor, N., Mischel, W., & Schwartz, J. C. (1982). A prototype analysis of psychological situations. *Cognitive Psychology, 14,* 45–77.

Cantor, N., Smith, E. E., French, R. D., & Mezzich, J. (1980). Psychiatric diagnosis as prototype categorization. *Journal of Abnormal Psychology, 89,* 181–193.

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: MIT Press.

Carroll, J. M. (1985). *What's in a name?* New York: W. H. Freeman.

Clark, E. (1983). Meaning and concepts. In J. H. Flavell & E. M. Markman (Eds.), *Handbook of child psychology: Volume III. Cognitive development* (pp. 787–840). New York: Wiley.

Cohen, B., & Murphy, G. L. (1984). Models of concepts. *Cognitive Science, 8,* 27–58.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8,* 240–247.

Donnellan, K. S. (1977). Reference and definite descriptions. In S. P. Schwartz (Ed.), *Naming, necessity, and natural kinds* (pp. 42–65). Ithaca, NY: Cornell University Press. (Reprinted from *Philosophical View,* 1966, *75,* 281–304)

Dupré, J. (1981). Natural kinds and biological taxa. *Philosophical Review, 90,* 66–90.

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

Estes, W. K. (1986b). Memory storage and retrieval processes in category learning. *Journal of Experimental Psychology: General, 115,* 155–174.

Estes, W. K. (1988). Toward a framework for combining connectionist and symbol-processing models. *Journal of Memory and Language, 27,* 196–212.

Estes, W. K., Campbell, J. A., Hatsopoulos, N., & Hurwitz, J. B. (1989). Base-rate effects in category learning: A comparison of parallel network and memory storage–retrieval models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 556–571.

Evans, G. (1977). The causal theory of names. In S. P. Schwartz (Ed.), *Naming, necessity, and natural kinds* (pp. 192–215). Ithaca, NY: Cornell University Press. (Reprinted from *Aristotelian Society Supplementary Volume,* 1973, *47,* 187–208)

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.

Fiske, S. T., & Linville, P. W. (1980). What does the schema concept buy us? *Personality and Social Psychology Bulletin, 6,* 543–557.

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.

Fodor, J. A. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences, 3,* 63–73.

Fodor, J. A. (1981). The present status of the innateness controversy. In J. A. Fodor (Ed.), *Representations* (pp. 257–316). Cambridge, MA: MIT Press.

Fodor, J. A. (1983). *The modularity of mind.* Cambridge, MA: MIT Press.

Fodor, J. A., Garrett, M. F., Walker, E. C. T., & Parkes, C. H. (1980). Against definitions. *Cognition, 8,* 263–367.

Fodor, J. D. (1977). *Semantics: Theories of meaning in generative grammar.* New York: Crowell.

Ford, M. (1985). Review of Johnson-Laird's "Mental Models." *Language, 61,* 897–903.

Franks, J. J., & Bransford, J. D. (1971). Abstraction of visual patterns. *Journal of Experimental Psychology, 90,* 65–74.

Fried, L. S., & Holyoak, K. J. (1984). Induction of category distributions: 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 judgements. *Cognitive Psychology, 16,* 341–370.

Gelman, S. A. (1988a). Children's expectations concerning natural kind categories. *Human Development, 31,* 28–34.

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

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development, 58,* 1532–1541.

Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition, 38,* 213–244.

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

Glass, A. L., Holyoak, K. J., & Santa, J. L. (1979). *Cognition.* Reading, MA: Addison-Wesley.

Gleitman, L. R., Armstrong, S. L., & Gleitman, H. (1983). On doubting the concept "concept." In E. K. Scholnick (Ed.), *New trends in con-*

ceptual representation: Challenges to Piaget's theory? (pp. 87–110). Hillsdale, NJ: Erlbaum.

Gluck, M. A. (1991). Stimulus generalization and representation in adaptive network models of category learning. *Psychological Science, 2,* 50–55.

Gluck, M. A., & Bower, G. H. (1988a). Evaluating an adaptive network model of human learning. *Journal of Memory and Language, 27,* 166–195.

Gluck, M. A., & Bower, G. H. (1988b). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General, 117,* 227–247.

Goldstone, R. L., Medin, D. L., & Gentner, D. (1991). Relational similarity and the nonindependence of features in similarity judgments. *Cognitive Psychology, 23,* 222–262.

Goodman, N. (1972). Seven strictures on similarity. In N. Goodman (Ed.), *Problems and projects* (pp. 437–446). New York: Bobbs-Merrill.

Hampton, J. A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning and Verbal Behavior, 18,* 441–461.

Hampton, J. A. (1981). An investigation of the nature of abstract concepts. *Memory & Cognition, 9,* 149–156.

Hampton, J. A. (1982). A demonstration of intransitivity in natural categories. *Cognition, 12,* 151–164.

Hampton, J. A. (1987). Inheritance of attributes in natural concept conjunctions. *Memory & Cognition, 15,* 55–71.

Hampton, J. A. (1988). Overextension of conjunctive concepts: Evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 12–32.

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

Haugeland, J. (1985). *Artificial intelligence: The very idea.* Cambridge, MA: MIT Press.

Hayes-Roth, B., & Hayes-Roth, F. (1977). Concept learning and the recognition and classification of exemplars. *Journal of Verbal Learning and Verbal Behavior, 16,* 321–338.

Heider, E. R. (1972). Universals in color naming and memory. *Journal of Experimental Psychology, 93,* 10–20.

Hersh, H. M., & Caramazza, A. (1976). A fuzzy set approach to modifiers and vagueness in natural language. *Journal of Experimental Psychology: General, 105,* 254–276.

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

Hintzman, D. L., & Ludlam, G. (1980). Differential forgetting of prototypes and old instances: Simulation by an exemplar-based classification model. *Memory & Cognition, 8,* 378–382.

Homa, D., Burruel, L., & Field, D. (1987). The changing composition of abstracted categories under manipulations of decisional change, choice difficulty, and category size. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 401–412.

Homa, D., Cross, J., Cornell, D., Goldman, D., & Schwartz, S. (1973). Prototype abstraction and classification of new instances as a function of number of instances defining the prototype. *Journal of Experimental Psychology, 101,* 116–122.

Homa, D., & Cultice, J. (1984). Role of feedback, category size, and stimulus distortion on the acquisition and utilization of ill-defined categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 83–94.

Homa, D., Dunbar, S., & Nohre, L. (1991). Instance frequency, categorization, and the modulating effect of experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 444–458.

Homa, D., Rhoads, D., & Chambliss, D. (1979). Evolution of conceptual structure. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 11–23.

Homa, D., Sterling, S., & Trepel, L. (1981). Limitations of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 418–439.

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

Horowitz, L. M., Wright, J. C., Lowenstein, E., & Parad, H. W. (1981). The prototype as construct in abnormal psychology: 1. A method for deriving prototypes. *Journal of Abnormal Psychology, 90,* 568–574.

Hull, C. L. (1920). Quantitative aspects of the evolution of concepts. *Psychological Monographs, 28,* (1, whole No. 123).

Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science, 4,* 71–115.

Johnson-Laird, P. N. (1981). Mental models of meaning. In A. K. Joshi, B. L. Webber, & I. N. Sag (Eds.), *Elements of discourse in understanding* (pp. 106–126). New York: Cambridge University Press.

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N. (1987). The mental representation of the meaning of words. *Cognition, 25,* 189–211.

Johnson-Laird, P. N., Herrmann, D. J., & Chaffin, R. (1984). Only connections: A critique of semantic networks. *Psychological Bulletin, 96,* 292–315.

Katz, J. (1972). *Semantic theory.* New York: Harper & Row.

Katz, J., & Fodor, J. A. (1963). The structure of a semantic theory. *Language, 39,* 170–210.

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* (pp. 175–200). New York: Cambridge University Press.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development.* Cambridge, MA: MIT Press.

Kelley, D., & Krueger, J. (1984). The psychology of abstraction. *Journal for the Theory of Social Behavior, 14,* 43–67.

Kellogg, R. T. (1981). Feature frequency in concept learning: What is counted? *Memory & Cognition, 9,* 157–163.

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

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85,* 363–394.

Klatzky, R. L. (1980). *Human memory: Structure and processes* (2nd ed.). New York: W. H. Freeman.

Knapp, A. G., & Anderson, J. A. (1984). Theory of categorization based on distributed memory storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 616–637.

Komatsu, L. K. (1983). *The structure of the mental lexicon.* Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

Kornblith, H. (1980). Referring to artifacts. *Philosophical Review, 89,* 109–114.

Kripke (1972). Naming and necessity. In D. Davidson & G. Harman (Eds.), *Semantics of natural language* (2nd ed., pp. 253–355). Dordrecht, Holland: Reidel.

Kripke, S. (1979). Speaker's reference and semantic reference. In P. A. French, T. E. Uehling, Jr., & H. K. Wettstein (Eds.), *Contemporary perspectives in the philosophy of language* (pp. 6–27). Minneapolis: University of Minnesota Press.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99,* 22–44.

Kunda, Z., Miller, D. T., & Claire, T. (1990). Combining social concepts: The role of causal reasoning. *Cognitive Science, 14,* 551–577.

Labov, W. (1973). The boundaries of words and their meanings. In C.-J. N. Bailey & R. W. Shuy (Eds.), *New ways of analyzing variation in English* (pp. 340–373). Washington, DC: Georgetown University Press.

Lakoff, G. (1972). Hedges: A study in meaning criteria and the logic of fuzzy concepts. In P. M. Peranteau, J. N. Levi, & G. C. Phares (Eds.), *Papers from the eighth regional meeting, Chicago Linguistics Society* (pp. 183–228). Chicago: Chicago Linguistics Society.

Lakoff, G. (1987a). Cognitive models and prototype theory. In U. Neisser (Ed.), *Concepts and conceptual development* (pp. 63–100). New York: Cambridge University Press.

Lakoff, G. (1987b). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Landau, B. (1982). Will the real grandmother please stand up? The psychological reality of dual meaning representations. *Journal of Psycholinguistic Research, 11,* 47–62.

Lingle, J. H., Altom, M., & Medin, D. L. (1984). Of cabbages and kings: Assessing the extendability of natural object concept models to social things. In R. S. Wyer, T. K. Srull, & J. Hartwick (Eds.), *Handbook of social cognition* (pp. 71–117). Hillsdale, NJ: Erlbaum.

Losonsky, M. (1990). The nature of artifacts. *Philosophy, 65,* 81–88.

Macnamara, J. (1991). *The takeover of psychology of biology or the devaluation of reference in psychology.* Unpublished manuscript, McGill University, Department of Psychology, Montreal, Quebec, Canada.

Malt, B. C. (1989). An on-line investigation of prototype and exemplar strategies in classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 539–555.

Malt, B. C. (1990). Features and beliefs in the mental representation of categories. *Journal of Memory and Language, 29,* 289–315.

Malt, B. C., & Johnson, E. C. (1992). Do artifact concepts have cores? *Journal of Memory and Learning, 31,* 195–217.

Malt, B. C., & Smith, E. E. (1984). Correlated properties in natural categories. *Journal of Verbal Learning and Verbal Behavior, 23,* 250–269.

Mayer, J. D., & Bower, G. H. (1986). Learning and memory for personality prototypes. *Journal of Personality and Social Psychology, 51,* 473–492.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 109,* 320–353.

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

McCloskey, M. E., & Glucksberg, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology, 11,* 1–37.

McNamara, T. P., & Miller, D. L. (1989). Attribute of theories of meaning. *Psychological Bulletin, 106,* 355–376.

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. (1983). Structural principles in categorization. In T. J. Tighe & B. E. Shepp (Eds.), *Perception, cognition, & development: Interactional analyses* (pp. 203–230). Hillsdale, NJ: Erlbaum.

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

Medin, D. L. (1989). Concepts and conceptual structure. *American Psychologist, 44,* 1469–1481.

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., Altom, M. W., & Murphy, T. D. (1984). Given versus induced category representations: Use of prototype and exemplar information in classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 333–352.

Medin, D. L., Dewey, G. I., & Murphy, T. D. (1983). Relationships between item and category learning: Evidence that abstraction is not automatic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 607–625.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1990). Similarity involving attributes and relations: Judgments of similarity and difference are not inverses. *Psychological Science, 1,* 64–69.

Medin, D., & 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). 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. (1981). Strategies and classification learning. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 241–253.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology, 35,* 113–138.

Medin, D. L., & Wattenmaker, W. D. (1987). Category cohesiveness, theories, and cognitive archeology. In U. Neisser (Ed.), *Concepts and conceptual development* (pp. 25–62). New York: Cambridge University Press.

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology, 19,* 242–279.

Mervis, C. B., Catlin, J., & Rosch, E. (1976). Relationships among goodness-of-example, category norms, and word frequency. *Bulletin of the Psychonomic Society, 7,* 283–284.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32,* 89–115.

Metcalfe, J., & Fisher, R. P. (1986). The relation between recognition memory and classification learning. *Memory & Cognition, 14,* 164–173.

Michalski, R. S. (1989). Two-tiered concept meaning, inferential matching, and conceptual cohesiveness. In S. Vosniadou & A. Ortony (Eds.) *Similarity and analogical reasoning* (pp. 122–145). New York: Cambridge University Press.

Miller, G. A. (1978). Practical and lexical knowledge. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 305–319). Hillsdale, NJ: Erlbaum.

Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception.* Cambridge, MA: Harvard University Press.

Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision* (pp. 211–277). New York: McGraw-Hill.

Mitchell, T. M., Keller, R. M., & Kedar-Cabelli, T. (1986). Explanation-based generalization: A unifying view. *Machine Learning, 1,* 47–80.

Murphy, G. L. (1988). Comprehending complex concepts. *Cognitive Science, 12,* 529–562.

Murphy, G. L. (1990). Noun phrase interpretation and conceptual combination. *Journal of Memory and Language, 29*, 259–288.

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 & Cognition, 13*, 377–384.

Neimark, E. D. (1983). There is one classification system with a long developmental history. In E. K. Scholnick (Ed.), *New trends in conceptual representation: Challenges to Piaget's theory?* (pp. 111–127). Hillsdale, NJ: Erlbaum.

Neisser, U. (1975). *Cognition and reality.* San Francisco: W. H. Freeman.

Neisser, U. (1987). From direct perception to conceptual structure. In U. Neisser (Ed.), *Concepts and conceptual development* (pp. 11–24). Cambridge, England: Cambridge University Press.

Nickerson, R. (1981). Motivated retrieval from archival memory. In (G. H. Bower, Ed.) *Nebraska symposium of motivation* (Vol. 28, pp. 73–119). Lincoln: University of Nebraska Press.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, NJ: Prentice-Hall.

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 104–114.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General, 115*, 39–57.

Nosofsky, R. M. (1988a). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal 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.

Nosofsky, R. M. (1989). Further tests of an exemplar-similarity approach to relating identification and categorization. *Perception and Psychophysics, 45*, 279–290.

Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception and Performance, 17*, 3–27.

Nosofsky, R. M., Clark, S. E., & Shin, H. J. (1989). Rules and exemplars in categorization, identification, and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 282–304.

Oden, G. C. (1987). Concept, knowledge, and thought. *Annual Review of Psychology, 38*, 203–227.

Ogden, C. K., & Richards, I. A. (1956). *The meaning of meaning* (8th ed.). New York: Harcourt, Brace.

Omohundro, J. (1981). Recognition vs. classification of ill-defined category exemplars. *Memory & Cognition, 9*, 324–331.

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.

Osherson, D. N. (1978). Three conditions on conceptual naturalness. *Cognition, 6*, 263–289.

Osherson, D. N., & Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition, 9*, 35–58.

Osherson, D. N., & Smith, E. E. (1982). Gradedness and conceptual combination. *Cognition, 12*, 299–318.

Osherson, D. N., Smith, E. E., Wilkie, O., López, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97*, 185–200.

Palmer, F. R. (1981). *Semantics* (2nd ed.). Cambridge, England: Cambridge University Press.

Panman, O. (1982). Homonymy and polysemy. *Lingua, 58*, 105–136.

Pazzani, M. J. (1991). Influence of prior knowledge on concept acquisition: Experimental and computational results. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 416–432.

Piaget, J. (1926). *The language and thought of the child.* New York: Harcourt, Brace.

Posner, M. I., & Keele, S. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77*, 353–363.

Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology, 83*, 304–308.

Putnam, H. (1975a). Is semantics possible? In H. Putnam (Ed.), *Philosophical papers: Volume 2. Mind, language and reality* (pp. 139–152). Cambridge, England: Cambridge University Press.

Putnam, H. (1975b). The meaning of "meaning." In H. Putnam (Ed.), *Philosophical papers: Volume 2. Mind, language and reality* (pp. 215–271). Cambridge, England: Cambridge University Press.

Quine, W. v. O. (1953). Two dogmas of empiricism. In W. v. O. Quine (Ed.), *From a logical point of view* (pp. 20–46). Cambridge, MA: Harvard University Press.

Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology, 3*, 382–407.

Reed, S. K. (1978). Category vs. item learning: Implications for categorization models. *Memory & Cognition, 6*, 612–621.

Rey, G. (1983). Concepts and stereotypes. *Cognition, 15*, 237–262.

Rey, G. (1985). Concepts and conceptions: A reply to Smith, Medin and Rips. *Cognition, 19*, 297–303.

Richards, D. D. (1988). Dynamic concepts and functionality: The influence of multiple representations and environmental constraints on categorization. *Human Development, 31*, 11–19.

Richards, D. D., & Goldfarb, J. (1986). The episodic memory model of conceptual development: An integrative viewpoint. *Cognitive Development, 1*, 183–219.

Richards, D. D., Goldfarb, J., Richards, A. L., & Hassen, P. (1989). The role of the functionality rule in the categorization of well-defined concepts. *Journal of Experimental Child Psychology, 47*, 97–115.

Rips, L. J. (1975). Inductive judgements about natural categories. *Journal of Verbal Learning and Verbal Behavior, 14*, 665–681.

Rips, L. J. (1986). Mental muddles. In M. Brand & R. M. Harnish (Eds.), *Problems in the representation of knowledge and belief* (pp. 258–286). Tuscon: University of Arizona Press.

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

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior, 12*, 1–20.

Robbins, D., Barresi, J., Compton, P., Furst, A., Russo, M., & Smith, M. A. (1978). The genesis and use of exemplar vs. prototype knowledge in abstract category learning. *Memory & Cognition, 6*, 473–480.

Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 111–144). San Diego, CA: Academic Press.

Rosch, E. (1975a). Cognitive reference points. *Cognitive Psychology, 7*, 532–547.

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

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.

Rosch, E. (1983). Prototype classification and logical classification: The two systems. In E. K. Scholnick (Ed.), *New trends in conceptual representation: Challenges to Piaget's theory?* (pp. 73–86). Hillsdale, NJ: Erlbaum.

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

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382–439.

Rosch, E., Simpson, C., & Miller, R. S. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 491–502.

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

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33–58). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 99–135). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Schemata and sequential thought processes in PDP models. In J. L. McClelland, D. E. Rumelhart, & The PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models* (pp. 7–57). Cambridge, MA: MIT Press.

Salmon, N. U. (1981). *Reference and essence.* Princeton, NJ: Princeton University Press.

Schank, R. C. (1982). *Dynamic memory: A theory of reminding and learning in computers and people.* Cambridge, England: Cambridge University Press.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding.* Hillsdale, NJ: Erlbaum.

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

Schwartz, S. P. (1978). Putnam on artifacts. *Philosophical Review, 87,* 566–574.

Schwartz, S. P. (1979). Natural kind terms. *Cognition, 7,* 301–315.

Schwartz, S. P. (1980). Natural kinds and nominal kinds. *Mind, 89,* 182–195.

Shafir, E. B., Smith, E. E., & Osherson, D. N. (1990). Typicality and reasoning fallacies. *Memory & Cognition, 18,* 229–239.

Shanks, D. R. (1990). Connectionism and the learning of probablistic concepts. *Quarterly Journal of Experimental Psychology, 42A,* 209–237.

Shanks, D. R. (1991). Categorization by a connectionist network. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 433–443.

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology, 52,* 1061–1086.

Simon, H. A. (1981). *The sciences of the artificial* (2nd ed.). Cambridge, MA: MIT Press.

Smith, E. E. (1978). Theories of semantic memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 6, pp. 1–56). Hillsdale, NJ: Erlbaum.

Smith, E. E. (1988). Concepts and thought. In R. J. Sternberg & E. E. Smith (Eds.), *The psychology of human thought* (pp. 19–49). Cambridge, England: Cambridge University Press.

Smith, E. E. (1989). Concepts and induction. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 501–526). Cambridge, MA: MIT Press.

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

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

Smith, E. E., & Osherson, D. N. (1984). Conceptual combination with prototype concepts. *Cognitive Science, 8,* 337–361.

Smith, E. E., Osherson, D. N., Rips, L. J., & Keane, M. (1988). Combining prototypes: A selective modification model. *Cognitive Science, 12,* 485–527.

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

Smith, L. B. (1989). From global similarities to kinds of similarities: The construction of dimensions in development. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 146–178). Cambridge, England: Cambridge University Press.

Smoke, K. L. (1932). An objective study of concept formation. *Psychological Monographs, 42* (4, Whole No. 191).

Smolensky, P. (1986). Information processing in dynamical systems: Foundations of harmony theory. In J. L. McClelland, D. E. Rumelhart, & The PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations* (pp. 194–281). Cambridge, MA: MIT Press.

Strange, W., Keeney, T., Kessel, F., & Jenkins, J. (1970). Abstraction over time of prototypes from distortions of random dot patterns. *Journal of Experimental Psychology, 83,* 508–510.

Taylor, S. E., & Crocker, J. (1980). Schematic bases of social information processing. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario symposium* (Vol. 1, pp. 89–134). Hillsdale, NJ: Erlbaum.

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

Tversky, A., & Gati, I. (1978). Studies of similarity. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 81–98). Hillsdale, NJ: Erlbaum.

Ward, T. B., & Scott, J. (1987). Analytic and holistic modes of learning family-resemblance concepts. *Memory & Cognition, 15,* 42–54.

Ward, T. B., Vela, E., & Hass, S. D. (1990). Children and adults learn family-resemblance categories analytically. *Child Development, 61,* 593–605.

Wattenmaker, W. D. (1991a). Learning modes, feature correlations, and memory-based categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 908–923.

Wattenmaker, W. D. (1991b, November). *Learning modes, feature correlations, and postencoding analyses.* Poster presented at the meeting of the Psychonomic Society, San Francisco.

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. J. Hilton (Ed.), *Contemporary science and natural explanation* (pp. 204–241). New York: New York University Press.

Welker, R. L. (1982). Abstraction of themes from melodic variation. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 435–447.

Whittlesea, B. W. A. (1987). Preservation of specific experiences in the representation of general knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 3–17.

Winograd, T. (1975). Frame representations and the declarative–procedural controversy. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 185–210). San Diego, CA: Academic Press.

Wittgenstein, L. (1953). *Philosophical investigations.* New York: Macmillan.

Zadeh, L. A. (1982). A note on prototype theory and fuzzy sets. *Cognition, 12,* 291–297.

Received January 2, 1991
Revision received December 19, 1991
Accepted December 19, 1991 ■