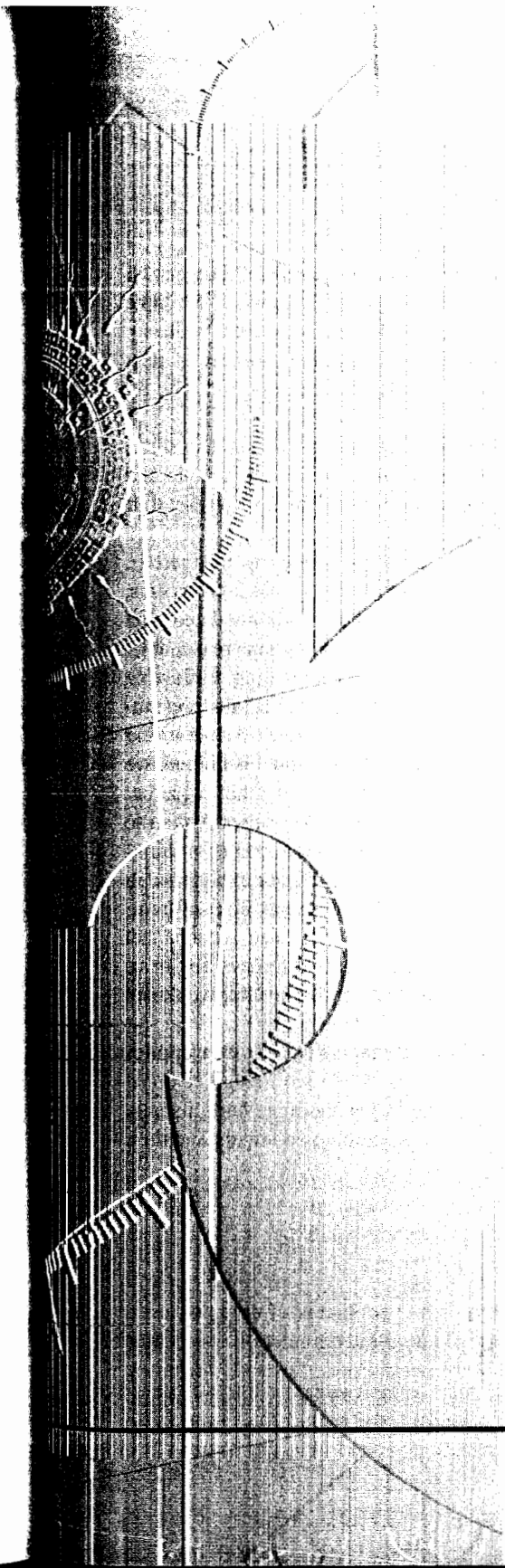


for commu-
998) are also
it source for
vides a fasci-
c (1974) is an
syntactic pro-
Dsherson and
offers a nice
auser (1996)
. Benson and
damage, and
source for on-



Chapter 10

Concepts and Categories: Representation and Use

Introduction

Why Categorize?
Computational Complexity
Functions of Concepts
Concepts and Misconceptions
Summary

Structure of Natural Object Categories

The Classical View
The Probabilistic View
Between-Category Structure
Does Similarity Explain Categorization?
Concepts as Organized by Theories
Putting Similarity in Its Place
Are There Kinds of Categories?
Summary

Use of Categories in Reasoning

Goals and Ad Hoc Categories
Conceptual Combination
Categories and Induction

Summary

Key Terms

Recommended Readings

*Mind and world in short have evolved together,
and in consequence are something of a mutual fit.*

—William James

INTRODUCTION



Why Categorize?

A basic cognitive function is to categorize. We have names for groups of things such as dogs, cats, cars, computers, birds, birthdays, and balloons. Likewise, we have names for properties (tall, tepid, tense) as well as actions (walk, waver, wallow). Furthermore, we can combine single words to create an unlimited number of new categories, like green-garbed grumpy golfers and camel-carted carpets. But why do we need categories?

The answer is that without categories we would be unable to make any sense of our experience or to profit from it. If each thing we encountered was unique and totally unlike anything else we had ever known, we would not know how to react to it or make any useful predictions about its properties. We would be literally lost in a sea of new experiences, helpless to employ any of our prior knowledge to navigate it. Imagine a clinical psychologist unwilling to form or use any diagnostic categories, who argues that every individual is unique and requires a totally individualized plan of treatment. The concept of unique treatments for every individual seems reasonable, even commendable, but when implemented it is completely self-defeating. Following Kendall (1975), suppose we have available treatment A and treatment B (and presumably more). What is our rationale for selecting which treatment to give? We are trying to predict which treatment will be more effective, but if each individual is unique, we have no basis whatsoever for making predictions! If we knew only that a patient seemed to be more similar to people who had responded well to treatment A than to those who had responded well to treatment B, we would have some rationale for thinking that treatment A would be more effective. Of course, to draw on our experience in this way is to categorize. Even before medicine had any sort of effective set of treatments, a major conceptual advance was made when people began to reject the idea that every instance of illness was unique. It seems inevitable, then, that to have any basis for providing (and even tailoring) a treatment, we need to categorize, even if we do not necessarily use formal diagnostic categories.

The need to categorize is not specific to clinical diagnosis but rather applies wherever relevant knowledge might be brought to bear. When we recognize some entity as a dog, our knowledge about dogs (e.g., that they sometimes bark, usually like to be petted, and so on) allows us to make predictions about and understand their actions. Categorization is pervasive.

Computational Complexity

It is easy to show that categorization quickly runs into problems of computational complexity. That is, we could categorize things in an unlimited number of ways, and we necessarily employ a minuscule subset of these possibilities. Suppose we have a set of n things (where n stands for some number). We can determine that the number of ways of assigning those things to categories increases very rapidly with n .

With two objects *a* and *b*, we can set up two categorization schemes: (*a*) (*b*) and (*ab*), where the parentheses define category boundaries. If we double the number of objects, we can have 15 distinct category structures: (*a*)(*b*)(*c*)(*d*), (*ab*)(*c*)(*d*), (*a*)(*bc*)(*d*), (*a*)(*b*)(*cd*), (*ac*)(*b*)(*d*), (*ad*)(*b*)(*c*), (*a*)(*bd*)(*c*), (*abc*)(*d*), (*a*)(*bcd*), (*acd*)(*b*), (*abd*)(*c*), (*ab*)(*cd*), (*ad*)(*bc*), (*ac*)(*bd*), and (*abcd*). By the time we get up to 10 objects, there are more than 100,000 possibilities!

The fact that there are so many possibilities makes it natural to ask why we have the categories we have rather than others. One possible answer is that our categories mirror the structure of the world. Perhaps the world comes organized into natural "clusters," and our concepts mirror those natural groupings or categories. An alternative possibility is that we have the categories we have because we are the sort of creatures we are. Another way of putting it is that the categories we have represent a solution for a set of problems (e.g., coping with ignorance, making predictions) and that perhaps we can better understand human concepts by asking what problems they address and what functions they serve (see Malt, 1995, for a cross-cultural perspective on these questions). In the following sections, we will review these functions, describe a body of research on category learning, and then return to the question of why we have the categories we have.

Functions of Concepts

So far, we have described concepts at a relatively informal level. When we try to be more precise, we see that concepts serve a number of distinct roles. In this discussion, we will often mention both *concepts* and *categories*. These terms have sometimes been used interchangeably in the literature on categorization. In this chapter, we use **concept** to refer to a mental representation and **category** to refer to the set of entities or examples picked out by the concept. Some researchers suggest that categories have independent existence in the world (independent of the organisms that conceive of them). We do not endorse this view because we believe that in many cases (if not all) categories are constructed by the human mind as it relates to the world. As we'll see, our categories have to be linked to the world if they are to prove useful.

Classification

A central function of concepts is to allow us to treat discriminably different things as equivalent. We use the word *things* here, because category members may be physical objects, living things, properties, actions, or even abstract ideas such as "democracy." Deciding that two (or more) items are members of the same category is the process of *classification*. As we will see, classification is one of the most widely studied functions of categories in psychology.

Understanding and Explanation

Classification allows intelligent organisms to break apart their experience into meaningful chunks and to construct an interpretation of it. A major facet of this

understanding is bringing old knowledge to bear on the current situation. For example, a person on a hike in the mountains who recognizes an animal as a rattlesnake will interpret the situation as dangerous. Concepts also support *explanations*; understanding why a friend reacted to a stick with alarm is explained with the knowledge that he or she initially classified it as a rattlesnake.

Prediction

A key aspect of classification is that it allows one to make *predictions* concerning the future, predictions that can be used to select plans and actions. For example, after we identify an animal as a rattlesnake, we can act to avoid it.

Reasoning

Concepts support *reasoning*. One does not need to store every fact and possibility if inferences can be derived from information that is stored. From the knowledge that all animals breathe, that reptiles are animals, and rattlesnakes are reptiles, one may reason (deductively) that rattlesnakes must also breathe, even though one may never have directly stored that fact. Furthermore, people can combine concepts to describe novel situations and to envision future states of affairs. You probably have never thought about or seen a *paper bee*. We asked a few people about this novel concept, and most of them came up with the idea that a paper bee is a bee made out of paper. Furthermore, everyone who arrives at this interpretation also envisions that such a bee would not be alive and could not breathe. To get an idea of how complex conceptual combination is, try to figure out the concept *paper committee*. Though it was natural to interpret paper bee as a bee *made of* paper, it seems more natural to interpret paper committee as a committee *concerned with* paper (or perhaps one that *exists* on paper, but not in reality) and certainly not one made out of paper.

Communication

To the extent that people share knowledge and index it in terms of the same categories, they will be able to communicate with each other. Communication allows learning on the basis of indirect experience. When an expert tells us to avoid sudden movements in the presence of rattlesnakes, we can follow this advice the very first time we are confronted with a rattlesnake.

Summary

One can readily see that concepts function in multiple ways and are essential to mental life. Understanding, explanation, and prediction are at the core of intelligent behavior. As we shall see, there is also a moral for researchers: the various conceptual functions may interact with and influence each other; for example, the reasoning and communication functions may affect categorization (e.g. Ross, 1997; Markman & Makin, 1998). Before turning to the question of how categories are structured, we focus on an important consequence of categorization.

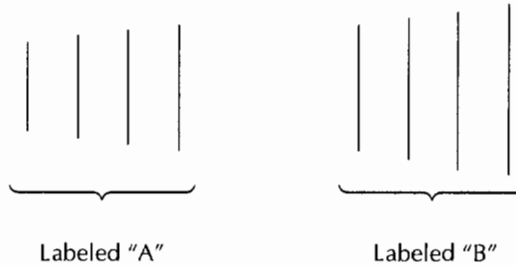
Concepts and Misconceptions

While categories are clearly crucial for human cognitive activity, our ability to use categories sometimes causes problems. We may form the wrong categories or categories that are too broad for our purposes. Furthermore, the very nature of our categories virtually ensures that important information may be lost or that inappropriate inferences will be made. On the one hand, our categories are a sign that some set of items should be treated equivalently in some way. On the other hand, we cannot treat the items from a common category exactly the same in every way. After all, both pit bulls and poodles are dogs, but one might not want to interact with them identically. Nonetheless, we may end up treating members of a category as being very similar simply because of their common category membership.

As an example, Tajfel and Wilkes (1963) showed people four short lines labeled "A" and four long lines labeled "B" and asked them to estimate the lengths of the lines. A control group was given the same task without the category labels (Figure 10.1). Relative to the control group, people given labels rated the short lines as more similar in length than did people in the control group and they also rated the short lines as more different from the long lines than people in the control group. In other words the labels made the examples within a category more similar and the differences across categories more distinctive.

Consequences of categorization are also evident in the important domain of social categorization. When people are brought into laboratories and divided into groups based on an arbitrary dimension (e.g., whether they overestimated or underestimated the number of dots in a picture), in-group favoritism results. What do we

Experimental condition



Control condition

Figure 10.1 Stimuli like those used in the Tajfel and Wilkes (1963) study.

mean by in-group favoritism? It is easiest to answer this by looking at an experiment, such as the one reported by Howard and Rothbart (1980). They created an arbitrary method for dividing people into groups. Howard and Rothbart showed people a page full of dots, and asked them to estimate how many dots were there. Then, they arbitrarily told people that they had overestimated or underestimated the number of dots. Further, they told people that there were groups of people who tended to overestimate and underestimate dots and that these people often had other characteristics in common. Then, each person was told favorable and unfavorable information about members of their group (the *in-group*) and the other group (the *out-group*). Despite the fact that this division into groups was arbitrary, people showed significantly better memory for negative behaviors of the out-group than for negative behaviors of the in-group. Thus, these categories influenced people's perceptions in a way that favored the in-group.

What about more socially relevant categories? Hirschfeld (1994, 1996) has shown that the development of children's understanding of race involves much more than a passive accumulation of information in the environment. Instead, it appears their thinking about race is organized by theories and beliefs about innate potential. For example, in one of his studies, Hirschfeld asked children to judge what the offspring of racially mixed couples would look like. Two groups of children (7- to 8-year-olds and 11- to 12-year-olds) from a suburban middle-class school (where there were few minority children) were shown pictures one at a time of four couples, consisting of a black male and a black female, a white male and a white female, a black male and a white female, or a white male and a black female. Each child then was shown pictures of three infants representing a white infant, a black infant, and an infant intermediate between the white and black infants in terms of skin color, hair color, and hair texture. Finally, each child was asked which of the infants was the child of the couple. The normatively correct answer for the mixed race couple is the infant depicted as intermediate.

All children judged that the black couple would have a black infant and the white couple a white infant. Of greatest interest are the judgments for the racially mixed couples. Younger children showed no clear preference among the three choices. Strikingly, however, older children overwhelmingly chose the black baby for both mixed race pairs. Furthermore, in a comparison condition involving skin and hair color of animals, older children make the normatively correct choice of intermediate color. These results suggest that by early adolescence children have learned the culturally dominant model of social potential—in our culture, children of black-white mixed couples are treated as black. (Regardless of whether one thinks this is sensible, it appears to be a fact.) Interestingly, when Hirschfeld gave the same task to children from an integrated inner-city school, both younger and older children tended to pick the infant depicted as intermediate. This pattern held for both children who identify as white and as black, suggesting that the results are driven by cultural environment rather than the children's racial status per se. In short, race is a salient social category that is susceptible to systematic misperceptions and cultural influences.

The results just described are somewhat sobering and certainly do not leave the impression that categorization is necessarily a good thing. The process of categorization leads only too naturally to stereotypes and misperceptions of other groups (see

STRUCTU



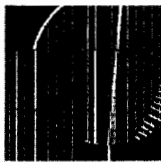
Devine, 1989; Tajfel, 1981, for examples). Perhaps knowledge of the fact of the human propensity for forming stereotypes can moderate or weaken our reliance on them. On a more optimistic note, there is evidence that stereotypes may yield to more concrete, specific information. Locksley, Borgida, Brekke, and Hepburn (1980) asked people to rate the assertiveness of target individuals who were described in one of three ways: (1) name only (conveying gender information), (2) name plus a descriptive paragraph irrelevant to assertiveness, or (3) name plus a descriptive paragraph relevant to assertiveness. In the first two conditions, male targets were rated as more assertive than female targets, but in the third females were rated as being as assertive as males; that is, when relevant information was provided, it was used and gender information was not used. In some cases, it seems, stereotypes are treated as default values that are employed only when no other information is available. Of course, what would be most helpful is a general theory about the conditions under which perceivers attend to and process information at the level of individuals versus categories. Both social psychologists and cognitive psychologists are keenly interested in this issue (e.g. Barsalou, Huttenlocher & Lamberts, 1998; Fiske, Neuberg, Beattie, & Milberg, 1987; Fiske, Lin, & Neuberg, 1999).

In short, categorization is necessary for all the reasons listed at the beginning of this chapter. But the benefits of categorization may come at some cost. Specifically, categorizing a set of objects can lead us to treat the members of the category as more similar to each other than they really are. In the case of social categories the result may be stereotyping.

Summary

In this section we noted that it is necessary to categorize in order to access relevant knowledge and make appropriate predictions and inferences. There are virtually an unlimited number of different ways in which we may categorize things and an important question is why we have the categories we have and not other ones. Part of the answer to that question is likely to be based on the different functions that categories serve (i.e. classification, understanding and explanation, prediction, reasoning, and communication). Finally we noted that for all of its benefits, categorization can both exaggerate (between category) differences and inappropriately minimize (within category) differences. We now turn our attention to the structural underpinnings of categories, beginning with natural object categories such as bird, fish, and tree.

STRUCTURE OF NATURAL OBJECT CATEGORIES



Almost all theories about the structure of categories assume that, roughly speaking, similar things tend to belong to the same category and that dissimilar things tend to be in different categories. For example, robins and sparrows both belong to the category *bird* and are more similar to each other than they are to squirrels or pumpkins. *Similarity* is a pretty vague term, but most commonly it is defined in terms of shared

properties or attributes. If you are asked to justify why you think robins and sparrows are more similar to each other than either is to squirrels you are likely to appeal to specific attributes or properties. For example, you might note that robins and sparrows are similar in that they are living, animate, have feathers, wings, and hollow bones, and can sing, fly, build nests, and lay eggs. Squirrels have only some of these properties, and pumpkins have fewer still. Although alternative theories assume concepts are structured in terms of shared properties, theories differ greatly in their organizational principles. Let's take a look at the main views concerning category structure.

The Classical View

The **classical view** assumes that concepts have defining features that act like criteria for determining category membership. For example, a triangle is a closed geometric form of three sides with the sum of the interior angles equaling 180 degrees. Each of these properties is necessary for an entity to be a triangle, and together these properties are sufficient to define *triangle*. According to the classical view, concepts have rigid boundaries in that a given example either does or does not meet the definition. All members of a category are equally good examples of it, and learning involves discovering these defining features.

Most of us have the initial intuition that our concepts conform to the classical view and have defining features. But think a bit more about a concept like chair or furniture. What makes a chair a chair? We might start by saying that a chair is something an individual can sit on, but that definition does not exclude sofas, benches, or even rocks. Next one might add the proviso that chairs must have four legs, but that would exclude beanbag chairs. The more we think about it, the trickier it becomes. For example, we might start to worry about the difference between stools and chairs and ultimately decide either that we do not know exactly what a chair is or that we may be unable to describe it.

There has, in fact, been a fair amount of research done on people's knowledge about object categories like bird, chair, and furniture. Not only do people fail to come up with defining features but also they do not necessarily agree with each other (or even with themselves when asked at different times) on whether something is an example of a category (Bellezza, 1984; McCloskey & Glucksberg, 1978). For example, is a rug considered furniture? A parquet floor? A telephone?

Philosophers and scientists also have worried about whether naturally occurring things like plants and animals (so-called natural kinds) have defining features. And the current consensus is that most natural concepts do not fit the classical view. Even the concept of species is not well defined (see Sokal, 1974). For example, a species might be defined as an interbreeding population. But in some species, the males of one group are fertile with the females of the other, but the females of the first group are not fertile with the males of the other. Although some concepts like triangle may be well defined, many concepts do not appear to be, and for this reason cognitive psychologists have pretty much abandoned the classical view.

The Probabilistic View

The major alternative to the classical view is the **probabilistic view**. It argues that concepts are organized around properties that may be typical or characteristic of category members. For example, most people's concept of bird may include the properties of building nests, flying, and having hollow bones, even though not all birds have these properties (e.g., ostriches, penguins).

The probabilistic view has major implications for how we think about categories. First, if categories are organized around characteristic properties, some members may have more of these properties than other members. In this sense, some members may be better examples or more typical of a concept than others, as we noted in Chapter 6. For example, people judge a robin to be a better example of a bird than an ostrich is and can answer category membership questions more rapidly for good examples than for poor examples (see, e.g., E. E. Smith, Shoben, & Rips, 1974; Rips, 1989; Medin & Heit, 1999, for reviews). A second implication is that category boundaries may be fuzzy. Nonmembers of a category may have almost as many characteristic properties of a category as do certain members. For example, whales have a lot of the characteristic properties of fish, and yet they are mammals. Third, learning about a category cannot be equated with determining what the defining features are because there may not be any.

Features and Typicality

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

Although typicality effects are robust and problematic for the classical view, the underlying basis for typicality effects may vary with both the kind of category being studied and with the population being studied. Barsalou (1985) showed that the internal structure of taxonomic categories is based primarily on the central tendency (or the average member) of a category. In contrast, the internal structure of goal-derived categories such as "things to wear in the snow" is determined by some ideal (or the best possible member) associated with the category. The best example of snow clothing, a down jacket, was not the example that was most like other category

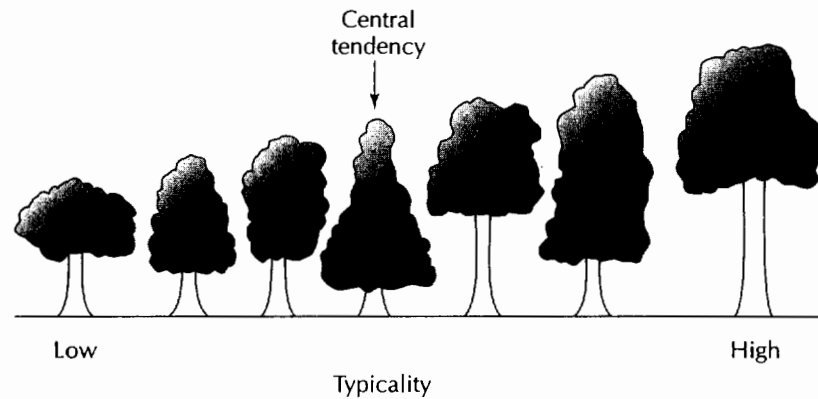


Figure 10.2

Typicality and central tendency illustrated for different kinds of trees varying in height. If typicality is based on central tendency, then trees of average height should be rated as the most typical. Tree experts, however, rate the tallest trees as most typical.

members; instead it was the example with the maximum value of the goal-related dimension of providing warmth.

One might hypothesize that ideals will only come into play when the category of interest lacks the natural similarity structure that characterizes common taxonomic categories such as bird, fish, and tree. However, recent evidence undermines this conjecture. Lynch, Coley, and Medin (2000) found that, for tree experts, the internal structure of the category *tree* was organized around the ideals of weediness and height (see Figure 10.2). The best examples of *tree* were not trees of average height but trees of extraordinary height (and free of “weedy” characteristics like having weak limbs, growing where they aren’t wanted, and being susceptible to disease).

It might occur to you that “best example” isn’t the same thing as “most typical.” However, Lynch et al. used exactly the same instructions employed by Rosch and Mervis (1975) in their original investigation of typicality effects. Moreover, Lynch et al. also ran undergraduate participants, and they showed no effects of ideals (their responses were mainly based on familiarity). In short, the differences in goodness of example effects appear to depend on expertise, not the wording of the instructions.

Atran (1998) reports a similar finding in studies comparing goodness of example effects for the category *bird* among University of Michigan undergraduates and Itzá Maya adults living in the rainforests of Guatemala. Undergraduates based typicality on overall similarity (central tendency) just as Rosch and Mervis (1975) had observed. The Itzá Maya, in contrast, based typicality on ideals—the best example of bird was the wild turkey which is culturally significant, prized for its meat, and strikingly beautiful. In short, ideals may play an important role in the internal organization of categories. There remains the question of why undergraduates differed from the Itzá Maya in the basis for their judgments. The fact that Lynch et al. (2000) found that U.S. tree experts based typicality on ideals suggests that it’s not just that the Itzá have a different notion of what typicality means. One speculation is that the

Figure 10.3



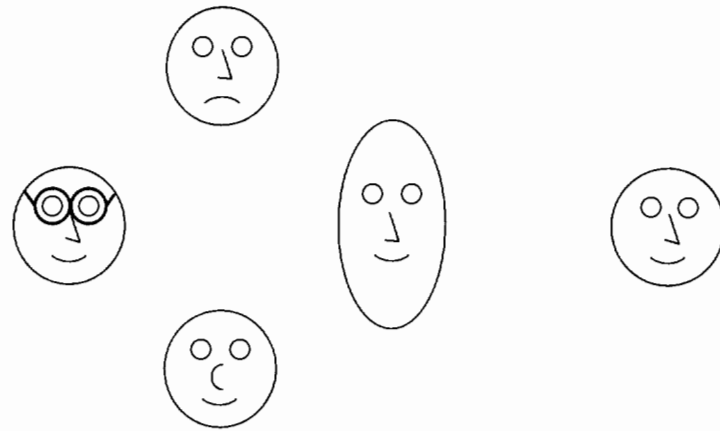
High

varying in height.
should be rated
typical.

The goal-related

When the category
common taxonomic
features, the internal
structure and
fuzziness and
of average height
features like having
features (like disease).
such as "most typ-
ical" employed by Rosch
features. Moreover,
effects of ideals
differences in good-
wording of the

the goodness of
of undergraduates
of undergraduates based
on Mervis (1975)
—the best exam-
ple for its meat, and
the internal orga-
nization of undergraduates differed
from Lynch et al. (2000)
but it's not just that
the relation is that the



Category Members

Prototype

Figure 10.3

Example of family resemblance. Each of the category members shares features with the other category members, but none of them are identical. The prototype on the right is the central tendency of these category members.

internal structure of categories is determined not only by classification processes but also by other conceptual functions (e.g., those mentioned at the beginning of this chapter). Undergraduates may know little about birds or trees and their interactions with them may not go beyond categorization (if that). Tree experts and Itzá Maya presumably have extended and complex interactions with trees and birds, respectively, and ideals have more opportunity to come into play. Perhaps similarity is a useful way to structure concepts used primarily for classification whereas ideals may be better for concepts used for a broader range of conceptual functions. (See Ross, 1997; Markman, Yamauchi, & Makin, 1997; and Solomon, Medin, & Lynch, 1999, for different ways of developing this argument.)

Mental Representations of Fuzzy Categories

If categories are not represented in terms of definitions, what form do our mental representations take? The term *probabilistic view* seems to imply that people organize categories via statistical reasoning. Actually, there is a more natural interpretation of fuzzy categories that has been referred to as a **family resemblance** principle. Many people in an extended family might share features such as a distinctive chin, certain expressions, or a high forehead. For example, Figure 10.3 sketches a hypothetical family. In this example, the faces all share many properties, though no two faces are identical. The general idea is that category members resemble each other in the way that family members do.

A simple summary representation for such a family resemblance structure would be an example that possessed all the characteristic features of a category. The best example is referred to as the **prototype**. For example, the prototype (based on

usual or modal values) for the faces of the category members in Figure 10.3 is shown on the right of the figure.

In a prototype model of categorization, classifying a new example is done by comparing the new item to the prototype. If the candidate example is similar enough to the prototype for a category, it is classified as a member of that category. The general notion is that, based on experience with examples of a category, people abstract out the central tendency or prototype that becomes the summary mental representation for the category. Note that no individual category member need have all the properties that are represented in the prototype. In that sense, a prototype is like a stereotype, which also may be true of no individual.

Laboratory studies of categorization using artificially constructed categories have been used to evaluate prototype theory. In these studies the experimenter creates the stimuli using salient properties and then assigns the stimuli to categories to create different kinds of category structures. For example, the stimuli might be geometric shapes varying in shape, size, and color, and fuzzy categories can be constructed by making sure that there is no simple rule that determines category membership. For example, one category might have examples that are usually large, usually red, and usually circular but there would be exceptions to each of these generalizations.

Normally many variables relevant to human category learning tend to be correlated with each other, and it is hard to determine which variables really are important. For instance, typical examples tend to be more frequent, and people's ability to classify typical examples quickly may be related to typicality, frequency, or both factors. The general rationale for laboratory studies with artificially created categories is that one can isolate some variable or set of variables of interest by breaking these natural correlations. Experiments with artificial categories reveal a number of salient phenomena associated with fuzzy categories, and several of these are consistent with prototype theories. For example, one observes typicality effects in learning and on transfer tests using both correctness and reaction time as the dependent measure (see, e.g., Rosch & Mervis, 1975). A striking result, readily obtained, is that the prototype for a category may be classified more accurately during transfer tests than are the previously seen examples that were used during original category learning (see, e.g., Homa & Vosburgh, 1976; Medin & Schaffer, 1978; Peterson, Meagher, Chait, & Gillie, 1973).

Typicality effects and excellent classification of prototypes are consistent with the idea that people are learning these fuzzy categories by forming prototypes. More detailed analyses, however, show problems with prototypes as mental representations. As we noted earlier, prototype theory implies that the only information abstracted from categories is the central tendency. A prototype representation discards information concerning category size, the variability of the examples, and correlations of attributes. The evidence suggests that people can use all three of these types of information (Estes, 1986; Flannagan, Fried, & Holyoak, 1986; Fried & Holyoak, 1984; Medin, Altom, Edelson, & Freko, 1982; Medin & Schaffer, 1978).

These same issues arise for real-world categories. For example, people seem to know that small birds are more likely to sing than large birds—a prototype does not capture this awareness of correlational information (Malt & Smith, 1983). In

In addition, how well an item belongs to a concept depends on the context in which it is presented (Roth & Shoben, 1983). For example, a harmonica is a typical musical instrument in the context of a campfire but atypical in the context of a concert hall. People seem to adjust their expectations in a manner sensitive to different settings. In short, prototype representations seem to discard too much information that can be shown to be relevant to human categorization.

Let's look at one further problem of prototype models in detail: predictions concerning category learning. To our knowledge, every model for category learning has some constraints or biases associated with it in the sense of predicting that some kinds of classification problems should be easier to master than others. One way to evaluate alternative learning models is to see if the problems they predict should be easy or difficult to acquire are, in fact, easier or more difficult for people to learn.

One constraint of interest is known as **linear separability**. Categories are linearly separable if one can categorize the examples perfectly by adding up the evidence from individual features. It is easiest to see what this means with a concrete example. Prototype models imply that categories must be linearly separable to be learnable. One way to conceptualize the process of classifying examples on the basis of similarity to prototype is that it involves a summing of evidence against a criterion. For example, if an instance has enough features characteristic of birds, it will be classified as a bird. More technically, there must be some weighted, additive combination of properties (some similarity function) that will be higher for all category members than for any nonmembers. If this is true, then the categories are linearly separable. All bird examples must be more similar to the bird prototype than to alternative prototypes, and all nonbirds must be more similar to their respective prototypes than to the bird prototype. If a bat were more similar to the bird prototype than to the mammal prototype, it would be incorrectly classified and the categories would not be linearly separable.

Figure 10.4 gives a more intuitive description of linear separability for examples that have values on two dimensions. In each graph category, members in

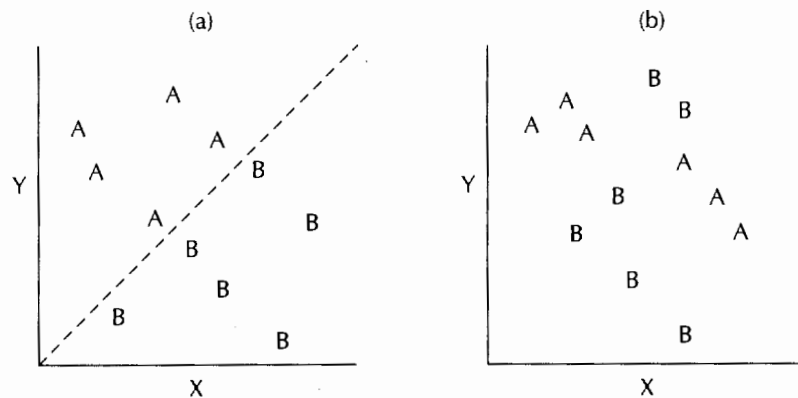


Figure 10.4 Two-dimensional example of two categories that are linearly separable (a) and are not linearly separable (b).

categories A and B are denoted respectively by the letters *A* and *B*. The position of each letter corresponds to its value on the X and Y dimensions (if it helps to be more concrete, think of these dimensions as the size and ferocity of animals). The categories are linearly separable if there is a straight line that perfectly partitions the categories (Figure 10.4a). If no straight line will partition the categories (Figure 10.4b), there is no way to construct prototypes such that all examples are closer to their own category prototype than to the prototype for the contrasting category. A number of neural network learning models favor categories that are linearly separable because, in effect, they add up the evidence favoring a classification decision (Minsky & Papert, 1988).

If linear separability acts as a constraint on human categorization, people should find it easier to learn categories that are linearly separable than categories that are not linearly separable. Studies employing a variety of stimulus materials, categories, subject populations, and instructions have failed to find any evidence that whether the categories are linearly separable influences the ease with which people learn categories (see, e.g., Kemler-Nelson, 1984; Medin & Schwanenflugel, 1981; see Smith, Murray, & Minda, 1997, for a dissenting opinion). This generalization has one striking exception: For social categories, there is strong evidence that linear separability does matter (Wattenmaker, 1995).

Wattenmaker suggests that social categories may be more compatible with a summing of evidence because (1) people may have a set of schemas or thematic structures that facilitate the integration of information across dimensions (e.g., the characteristic features knowledgeable, competent, hard working, and composed can readily be integrated if the category *scientist* is activated), and (2) people show considerable flexibility in relating features of examples to more abstract underlying properties (e.g., the characteristic features relaxed, thoughtful, and friendly could be seen as consistent with the abstract trait *intelligent*, but they may also be seen as consistent with the trait *kind*). Even with this salient exception of social categories, the sort of domain-general linear separability constraint implied by prototype models does not appear to hold. Overall, then, prototype models appear to have a number of serious limitations. An alternative approach, which is also consistent with the probabilistic view, assumes that much more information about specific examples is preserved. This approach appropriately falls under the general heading of exemplar theories.

Exemplar Theories

Exemplar theories provide an alternative way of representing probabilistic or fuzzy categories. **Exemplar models** assume that people initially learn some examples of different concepts and then classify a new instance on the basis of how similar it is to the previously learned examples. The idea is that a new example reminds the person of similar old examples and that people assume that similar items will belong to the same category. For example, you might classify one animal as a rodent because it reminds you of a mouse (which you know is a rodent) but classify some other animal as a rodent because it reminds you of a squirrel (which you also know is a rodent). As another example, suppose you are asked whether large birds are more or

less likely to fly than small birds. You probably will answer "less likely," based on retrieving examples from memory and noting that the only nonflying birds you can think of are large (e.g., penguin, ostrich).

The exemplar models that have received the most attention (Brooks, 1978; Hintzman, 1986; Medin & Schaffer, 1978; Nosofsky, 1986) assume that examples that are most similar to the item to be classified have the greatest influence on categorization. This could arise because the likelihood of retrieving an example from memory depends on its similarity to the item. The idea that retrieval is similarity-based and context-sensitive is in accord with much of the memory literature (see, e.g., Tulving, 1983; Chapter 5). Surprisingly, exemplar models can even account for the observation that the prototype may be more accurately classified on a transfer test than examples seen during original learning. The reason is as follows: The prototype will tend to be similar to many examples from its own category and not very similar to examples from alternative categories. Therefore, the prototype should reliably remind the learner of examples from the correct category. In contrast, some of the examples seen during training may not be highly similar to some of the other examples from their own category and may actually be similar to examples from other categories. In this case, the example may remind the learner of examples from alternative categories. It is important to bear in mind that exemplar models do not assume that people are necessarily able to retrieve individual examples one at a time without confusing them. Instead, the idea is that a test example will tend to activate a number of similar stored representations.

Quite a few experiments have contrasted the predictions of exemplar and prototype models. In head-to-head competition, exemplar models have been substantially more successful than prototype models (Barsalou & Medin, 1986; Estes, 1986, 1994; Lamberts, 1995; Medin & Coley, 1998; Nosofsky, 1988a, 1988b, 1991; Nosofsky & Palmeri, 1997) even for natural language categories (Storms, De Boeck, & Ruts, 2000). See Homa, 1984, and Smith & Minda, 1998, for opposing views.

Why should exemplar models fare better than prototype models? One of the main functions of classification is to allow one to make inferences and predictions on the basis of partial information (see J. R. Anderson, 1990a, 1990b). Here we are using classification loosely to refer to any means by which prior (relevant) knowledge is brought to bear, ranging from a formal classification rule to an idiosyncratic reminding of a previous case (which, of course, is in the spirit of exemplar models; see also Kolodner, 1993). Relative to prototype models, exemplar models tend to be conservative about discarding information that facilitates predictions. For instance, sensitivity to correlations of properties within a category enables finer predictions: From noting that a bird is large, one can predict that it cannot sing. In short, exemplar models better allow predictions and inferences than prototype models.

Exemplar models do seem to preserve more information than prototype models. Because the prototype is an average exemplar, it loses information about the specific category members that were seen. Although the exemplar model stores the category members, that does not mean that it will ever be possible to retrieve that single exemplar, because activating one exemplar is likely to activate others that are similar to it. Generally speaking, however, exemplar models preserve more information

AN APPLICATION**Exemplar Similarity in Medical Diagnosis**

Medical diagnosis is an important categorization task. Therefore, one ought to expect that models of categorization would be relevant to medicine and that questions about prototypes versus exemplars might apply to medicine. Brooks, Norman, and Allen (1991) demonstrated that exemplar similarity influences medical diagnosis and not just performance on tasks set up in psychology laboratories. They ran experiments with medical residents and general practitioners involving diagnoses of dermatological (skin) problems. The materials they used were color slides of dermatological lesions from the slide libraries of two practicing and teaching dermatologists. Slides were selected from a range of diagnostic categories to construct examples for a study phase (30 slides) and a test phase (60 slides). The test phase slides included some examples from the study phase, some selected to be very similar to a study example, and some that were not very similar but involved the same category. Not surprisingly, the physicians were more accurate on old examples that they had studied than for new examples. Of greatest interest is the fact that physicians were

(Continued)

than prototype models—information that people seem to be able to use. This does not prove that exemplar models are correct, but it suggests that learning depends more on examples than prototype models have assumed (Medin & Ross, 1989).

In short, exemplar models appear to have a number of effective characteristics, and future generations of models will likely embody these characteristics, directly or indirectly. For example, recent models of category learning that combine searching for rules with memory for examples have proven successful, at least for artificial categories (e.g. Nosofsky, Palmeri, & McKinley, 1994; Erickson & Kruschke, 1998; see also Ashby, Alfonso-Reese, Turken, & Waldron, 1998, for another dual-process approach to categorization). Instance-based ideas have also been incorporated into artificial intelligence categorization models (see, e.g., Kibler & Aha, 1987; Stanfill & Waltz, 1986). Finally, we note that there have been formal proofs that classifying on the basis of the single most similar example (a “nearest neighbor” principle) is not far from the optimal classification (Cover & Hart, 1967).

Although exemplar models have the virtue that they discard very little information that is potentially relevant to later judgments, the exemplar notion is incomplete as a theory of categories. Items are assumed to be placed in the category with the exemplars they are most similar to, but exemplar theories do not explain how concepts are created in the first place; that is, there are no constraints on what can be

AN APPLICATION

Exemplar Similarity in Medical Diagnosis (*Continued*)

more accurate on a new example highly similar to an old example than they were on a new, less similar example involving the same category.

This finding is surprising because both the residents and the general practitioners had considerable prior experience with dermatological diagnosis. One estimate is that about 6% of the cases general practitioners see involve dermatology. The physicians in the study by Brooks et al. had between 5 and 49 years of postgraduate experience. Nonetheless, the effect of specific example similarity did not diminish as a function of experience. It also appeared on both an immediate test and after a two-week delay between the study phase and the test phase. The effect was on the order of 10% to 20%, which is not only statistically reliable but clinically significant. This result is important because it suggests that if you need to see a doctor, the diagnosis may depend partially on the sorts of cases the doctor has seen recently. In summary, categorization processes studied in the laboratory are relevant to real world applications and a better understanding of these processes may have important application to medical diagnosis and training.

a concept. Furthermore, the only explanation for why a new example should be placed into a category is "because it is similar to an old example"—a very limited form of explanation. In addition, these models focus on how people learn to classify objects to the exclusion of other functions like prediction, reasoning, or communication. As we shall see, these three problems are also true for probabilistic view theories.

Summary

The proposals about types of category representations are summarized in Figure 10.5. Despite the strong intuition that categories are organized around defining features, the classical view of concepts does not account for a variety of observations (e.g., fuzzy categories, typicality effects) and the consensus favors the idea that concepts are structured more probabilistically. For this reason, various probabilistic views of categorization have been developed. The two most prominent theories of this type are the prototype and exemplar models. The prototype model assumes that category representations consist of a central tendency or average member of a category. The exemplar model assumes that people store individual exemplars of categories, and then classify new instances with respect to their similarity to these stored exemplars. In general, it appears that people's categorization shows a context-sensitivity that is

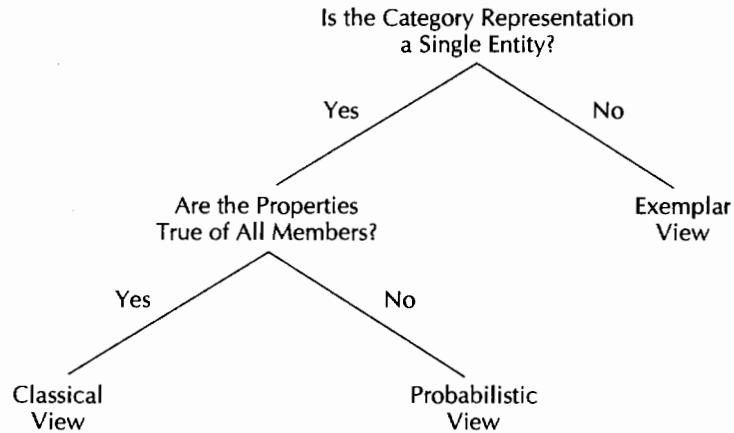


Figure 10.5 Summary of the types of category representations.

better described by exemplar models than prototype models. So far, however, we have focused mainly on the categorization functions of concepts. It is not clear how probabilistic view models will address other conceptual functions such as conceptual combination (see Kamp & Partee, 1995; Osherson & Smith, 1982, 1997; Fodor & Lepore, 1996, for discussion).

The discussion to this point has focused exclusively on the way information about individual categories is represented (or *within-category structure*). It is almost meaningless to talk about within-category structure without also worrying about how different categories are organized with respect to each other (or *between-category structure*). For example, when we say that examples within a category tend to be similar to each other we mean that they are more similar than pairs of examples that come from different categories. We turn now to some ideas concerning between-category structure before worrying again about why we have the categories we have and not others.

Between-Category Structure

Most things have membership in numerous categories. In this section we consider two types of between-category structure: a hierarchical, taxonomic structure and structures such as social categories that provide multiple, overlapping contrasts with, at best, a weak hierarchical structure.

Hierarchical Structure and the Basic Level

Often categories are organized *taxonomically* at different levels of abstraction. At each level, an example belongs to one of a set of mutually exclusive categories (as

Figure 10.6

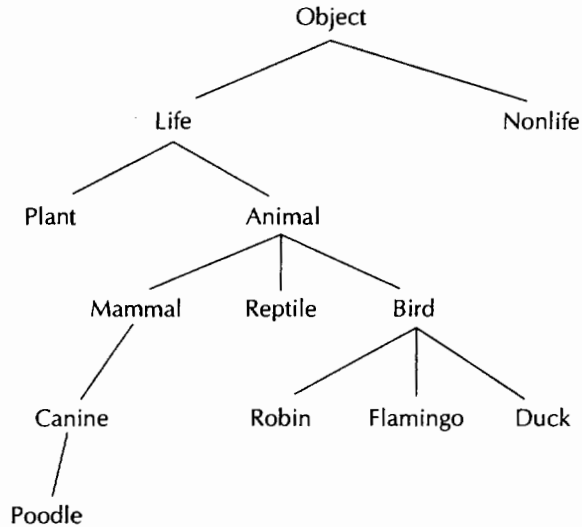


Figure 10.6 Part of a simple taxonomic structure.

we saw in our discussion of models of semantic knowledge in Chapter 6). A poodle is also a canine, a mammal, an animal, a living thing, and an object (Figure 10.6). A poodle is a mammal and not a reptile, an animal and not a plant, and so on.

An important observation about categories is that people are not confused about which label to use when asked about an object. For example, if you were looking at a four-legged furry animal that had just fetched a ball, you would be more likely to call it a *dog* than an *animal* or a *golden retriever*. This middle level of abstraction, which seems to provide the label that we would use as a default is called the **basic level**, and it appears to be psychologically privileged. For example, basic level concepts are the first to be learned, the natural level at which objects are named, and the highest level in which the instances all share the same parts and overall shape (Rosch, Mervis, Gray, Johnson, & Bayes Braem, 1976). The basic level resides at a middle level of abstraction. More abstract categories (like animal) are called *superordinates*, and more specific categories (like golden retriever) are called *subordinates*. There is some evidence that as one becomes more expert in an area, what was previously the subordinate level becomes the basic level. For example, dog might be a basic-level category for most people, but poodle might be at the basic-level for dog trainers (Tanaka & Taylor, 1991; see also Johnson & Mervis, 1997, 1998).

If we could understand just why the basic level is basic, we might gain further insight into what types of categories are natural or cohesive. One possibility is that the basic level is the highest level at which entities tend to share parts (B. Tversky & Hemenway, 1984; see Markman & Wisniewski, 1997, for a related idea). Another idea is that levels higher than the basic level have different purposes. For example, superordinate categories may serve to organize scenes (Murphy & Wisniewski, 1989; see

AN ENIGMA.

Will the True Basic Level Please Stand Up?

The idea of a basic level did not originate in cognitive psychology—it began with anthropological studies examining categorization in different cultures. The lion's share of this work was done in ethnobiology. The general question was whether people in different cultures organize biological categories in similar ways (that is, whether there are universal principles of biological categorization). The general answer proved to be "yes" but cross-cultural agreement in people's categorization of plants and animals appears to be especially strong at one level in taxonomic hierarchies, the basic level (Atran, 1990; Berlin, 1992; Berlin, Breedlove, & Raven, 1973; see Malt, 1995, for a review). But herein lies a yet-unresolved puzzle. The level that the cross-cultural studies of biological categorization suggest is basic corresponds more or less to the genus level (e.g., maple) in scientific taxonomy. Rosch et al., however, found that, for Berkeley undergraduates, examples of this level acted like subordinates. Rather than maple, oak, trout, and cardinal being basic, Rosch et al. (1976) found that tree, fish, and bird met their criteria for basicness.

Why do ethnobiological and psychological measures of the basic level disagree? One possibility is that the Berkeley undergraduates in Rosch's studies knew little about biological categories, especially relative to the people studied in cross-cultural investigations. In other words, maybe the difference is a difference in expertise (as in the findings of Tanaka and Taylor described earlier which suggest that the basic level may become more specific with expertise). A second possibility grows out of the fact that exactly comparable measures of basicness were not used; different measures may pick out different levels as privileged. Ethnobiological studies tend to use naming or linguistic criteria for basicness, whereas Rosch et al. relied more heavily on feature listings and perceptual tests. Interestingly, the clearest changes with expertise in the Tanaka and Taylor studies involved naming preferences. Until this puzzle is resolved we have to admit that we do not know exactly what makes the basic level basic.

More recently some direct cross-cultural comparisons have been done using comparable measures. Coley, Medin, and Atran, 1997 (see also Atran, Estin, Coley, & Medin, 1997), compared U.S. undergraduates
(Continued)

also Wisniewski, Imai, & Casey, 1996), and superordinate category members may tend to share common functions more than perceptual similarities (see, e.g., Rosch et al., 1976; Murphy & Smith, 1982).

AN ENIGMA

Will the True Basic Level Please Stand Up? (Continued)

and Itzá Maya adults of Guatemala using *inductive confidence* (reasoning from being told that one member has some novel property to all members having that property) as a measure of basicness. They assumed that if the basic level is the most abstract level at which category members share many properties, then inductive confidence should drop abruptly for categories above the basic level. For example, if U.S. undergraduates know that birds share many properties, then they might be confident that a property true of robins will be true of all birds. But since there are many fewer properties that all members of the category animals share, they presumably would not be at all confident that a property true of robins would be true of all animals. The Itzá Maya are members of a more traditional society and they rely on the rainforest for their livelihood. They might show a greater differentiation of the category bird (see larger differences among birds), and they might believe that a property true of a specific bird would not necessarily be true of all birds. This is the prediction one would make from the point of view that people in traditional societies tend to be biological experts, relative to U.S. undergraduates. Surprisingly, both the Itzá Maya and U.S. undergraduates consistently privileged the same level and this corresponded to the level of genus (robin, trout, and oak rather than bird, fish, and tree), consistent with expectations derived from anthropology.

These results raise a number of new puzzles that can only be answered by future research. Why do undergraduates show a different privileged level for induction versus other tasks such as speeded categorization or feature listing? One possibility draws on the distinction between knowledge and expectations. Undergraduates might be hard-pressed to list features that distinguish elms from oaks but they might nonetheless expect them to have lots of differences. It may be that the inductive confidence task gets at expectations rather than knowledge. Another question is how people from traditional cultures would perform on Rosch's perceptual tasks. Would they be fastest at the level of categorizing examples into bird, fish and tree or robin, trout, and oak? The Tanaka and Taylor studies with bird and dog experts suggest that the Itzá Maya might be fastest at the genus level, because the Itzá are almost certainly biological experts. Cross-cultural comparisons are needed to address this question.

Nonhierarchical Categories

Many of the categories that we use do not fit into a taxonomic hierarchy. This is particularly true in the domain of socially-relevant concepts. One can be a mother, a

biology—it
in different
y. The gen-
e biological
l principles
e "yes" but
nd animals
rchie, the
ven, 1973;
puzzle. The
suggest is
e) in scien-
undergrad-
han maple,
d that tree,

basic level
; in Rosch's
ative to the
, maybe the
Tanaka and
ay become
t of the fact
ed; different
logical stud-
reas Rosch
sts. Interest-
aylor studies
we have to
l basic.

have been
1, 1997 (see
l undergraduates
(Continued)

ry members may
(see, e.g., Rosch

psychologist, a Democrat, a golf player, and a Mexican-American all at the same time, and no single category is either superordinate or subordinate with respect to the others. If there is no clear hierarchy, then by definition there is no basic level. Given that people tend to use categories to understand their experience, one can think of the various categorization schemes as competing for attention. We already know from the memory literature that it is implausible to think that to understand some behavior, people access all the potentially relevant categories they have and reason with them. So what does determine which categories people access and use? Two factors that influence category access are the frequency and recency with which a category has been used. For example, in one study (Srull & Wyer, 1979, Experiment 1), people first performed a sentence construction task that was designed to activate concepts associated with hostility. Then, in an ostensibly separate experiment, the participants were asked to form impressions of people based on reading a description of behaviors. The behaviors were ambiguous with respect to hostility. Ratings of the descriptions with respect to hostility increased directly with the number of times hostility-related concepts had been activated on the presumably unrelated sentence construction task. The Srull and Wyer study illustrates the point that category accessibility has an important influence on how social information is encoded and interpreted (see also Smith, Fazio, & Cejka, 1996).

Is there a notion of privilege for nonhierarchical categories? Some social categories, such as those marking race, age, and gender, may be accessed automatically (e.g. Bargh, 1994; Greenwald & Banaji, 1995) independent of any intentions on the part of people. There is even some intriguing evidence that the activation of one social category leads to the inhibition of competing social categories (Macrae, Bodenhausen, & Milne, 1995). Although the structural differences between hierarchical and nonhierarchical categories are important, there is little or no work that has directly compared them.

Summary

The idea of a privileged or basic level is a very important principle of categorization. Originally it was interpreted as indicating that the world is organized into natural chunks that more or less impose themselves on minds. Evidence that the basic level changes with expertise suggests that privilege depends on the interaction of goals, activities, and experience with objects and events in the world.

So far all of our discussion of within- and between-category structure has explicitly or implicitly assumed that similarity relations determine categorization (the classical view can be seen as a special case where only defining features contribute to similarity). We turn now to a closer examination of similarity. As we shall see, the notion of similarity can be a bit slippery. It is a very important construct but one must be very careful about how it is used.

Does Similarity Explain Categorization?

Similarity is a very intuitive explanation for how people categorize—we put things into the same category because they are similar. But similarity is very difficult to pin

Figure 10.7

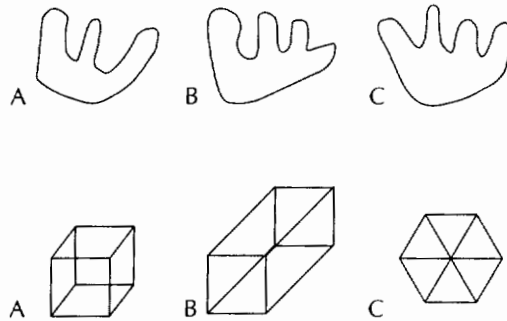


Figure 10.7

Sample of stimuli used by Medin, Goldstone, and Gentner (1993). For each set, people either saw A and B or B and C. The task was to list common and distinctive properties.

down. One problem with using similarity to define categories is that similarity is too variable. An example of the context-dependent nature of similarity is shown in Figure 10.7 (taken from Medin, Goldstone, & Gentner, 1993). People were either shown stimulus A and B together or stimulus B and C from sets of three like those shown in the figure. For each pair, people were asked to list common (shared) and distinctive features. Medin et al. (1993) found that the properties ascribed to stimulus B depended on whether it was paired with A or C. For example, for the top set, B was described as having three prongs when it was paired with A but four prongs when it was paired with C. For the bottom set, B was seen as three-dimensional when paired with A but two-dimensional when paired with C. Such observations suggest that similarity is heavily influenced by context.

Formal models of similarity allow similarity to be quite flexible. For example, Tversky's (1977) influential contrast model defines similarity as depending on common and distinctive features weighted for salience or importance. According to this model, similarity relationships depend heavily on the particular weights given to individual properties or features. A zebra and a barber pole could be seen as more similar than a zebra and a horse if the feature striped is given sufficient weight. This would not necessarily be a problem if the weights were stable. However, Tversky and others have convincingly shown that the relative weighting of a feature (as well as the relative importance of matching and mismatching features) varies with the stimulus context, experimental task (Gati & Tversky, 1984; Tversky, 1977), and probably even the concept under consideration (Ortony, Vondruska, Foss, & Jones, 1985). For example, a person from Maine and a person from Florida will seem less similar when they meet in Washington, DC, than when they meet in Tokyo.

Once we recognize that similarity is dynamic and depends on some (not well understood) processing principles, earlier work on the structure of fuzzy categories can be seen in a somewhat different light. Recall that the Rosch and Mervis (1975) studies asked subjects to list attributes or properties of examples and categories. It would be a mistake to assume that people had the ability to read and report their mental representations of concepts in a perfectly accurate manner. Indeed, Keil

(1979, 1981) pointed out that examples like robin and squirrel share many important properties that almost never show up in attribute listings (e.g., has a heart, breathes, sleeps, is an organism, is an object with boundaries, is a physical object, is a thing, can be thought about, and so on). In fact, Keil argued that knowledge about just these sorts of properties serves to organize children's conceptual and semantic development. For present purposes, the point is that attribute listings provide a biased sample of people's conceptual knowledge.

To take things a step further, one could argue that without constraints on what is to count as a feature, any two things may be arbitrarily similar or dissimilar. Thus, as Murphy and Medin (1985) suggested, the number of properties that plums and lawn mowers have in common could be infinite: Both weigh less than 1,000 kg, both are found on the earth, both are found in our solar system, both cannot hear well, both have parts, both are not worn by elephants, both are used by people, both can be dropped, and so on (see also Goodman, 1972; Watanabe, 1969). Now consider again the status of attribute listings. They represent a biased subset of stored or readily inferred knowledge. The correlation of attribute listings with typicality judgments is a product of such knowledge and a variety of processes that operate on it. Without a theory of that knowledge and those processes that use it, it simply is not clear what these correlations indicate about mental representations.

If similarity is so flexible, how can it provide the basis for determining categories? One possibility is that children are not as flexible about similarity as adults. Linda Smith (1989) has proposed that there is a developmental increase in the tendency of children to weight dimensions differentially, with the youngest children biased toward responding in terms of overall similarity. If flexibility arises only after most perceptual categories are learned, then similarity might be stable enough to support the initial development of many categories. A closely related idea is that children rapidly learn what types of similarity matter in particular contexts.

In part, the question may be just how flexible similarity is. For example, the object recognition theories reviewed in Chapter 4 certainly assume that similarity is stable enough to be useful. In particular, structural approaches to similarity such as geon theory may be more constrained than featural approaches (because it is the combination of geons and their relations that is important rather than some flexible weighting of individual features) and consequently may provide more stability (e.g., Medin et al., 1993; Goldstone, 1994a).

Even if similarity can be constrained, it still may not explain categorization. We believe that similarity is properly viewed as a general guideline or heuristic for categorization but that it does not provide the backbone of conceptual structure. Things that look alike do tend to belong to the same category. (But there are exceptions: Mannequins may look enough like people for us to confuse them with the real thing, but only briefly.) Furthermore, things that look alike superficially often tend to be alike in other, deeper ways (e.g., structure and function tend to be correlated). Overall similarity is a good but fallible guideline to category membership. When additional information (or a closer look) suggests that overall similarity is misleading, it is readily abandoned.

Figure 10.8



Figure 10.8 Sample triplet of stimuli from Gelman and Markman (1986). The task was to decide whether the bottom figure had a property true of the perceptually more similar choice (top right) or a property true of the example that belonged to the same category (top left).

A nice illustration of the fact that even young children are not strongly constrained by overall similarity comes from a set of experiments by Gelman and Markman (1986). They pitted category membership against perceptual similarity in an inductive reasoning task (an example is shown in Figure 10.8). Young children were first shown pictures of two animals and taught that different novel properties were true of them. Then they were asked which property was also true of a new (pictured) example that was perceptually similar to one alternative but belonged to the category of the other alternative, from which it differed perceptually. For example, children might be taught that a flamingo feeds its baby mashed-up food and that a bat feeds its baby milk, and then asked what a blackbird feeds its baby (see Figure 10.8). The blackbird was perceptually more similar to the bat than to the flamingo, but even 4-year-olds made inferences on the basis of category membership rather than similarity. That is, they thought that the blackbird would feed its baby mashed-up food. Therefore, even for young children, similarity acts as a general guideline that can be overridden by other forms of knowledge.

But similarity is not a notion to be lightly dismissed. Even if similarity is not some bedrock principle for *explaining* categorization it may nonetheless play an extremely important role. One way to summarize the upshot of current research is to say that similarity affects categorization *and* categorization affects similarity as well. There is increasing evidence that the features or building blocks of categories are not hard-wired into the nervous system but rather can be modified by experience (e.g. Gauthier & Tarr, 1997; Goldstone, 1994b; Oliva & Schyns, 1997; Schyns & Rodet, 1997; Schyns, Goldstone, & Thibaut, 1998). There is also abundant evidence that the relative weight given to different kinds of features varies with their relevance to a given categorization task. For example, in Chapter 4, we discussed that people can attend either to fine detail information or coarse information when processing an

image. Schyns and Oliva (1999) extended this finding by demonstrating that people could learn to preferentially attend to either a coarse spatial scale or a fine spatial scale in a speeded categorization task. These results suggest that categorization and (perceptual) similarity are closely intertwined, so much so that one could as well say that categorization causes similarity as the converse.

Summary

It does not seem that similarity, at least in the form that it takes in current theories, is going to be adequate to explain categorization. As we have seen, similarity may be the outcome or by-product rather than the cause of categorization. To use a rough analogy, winning basketball teams have in common scoring more points than their opponents, but one must turn to more basic principles to explain why they score more points. Similar things may share properties and be in the same category, but similarity may not explain *why* they are in the same category. We will explore the process by which similarity is determined in Chapter 11. Now, we will focus on the idea that concepts are organized around theories. In the next section, we will briefly summarize some of the current work on the role of knowledge structures and theories in categorization and then turn to a way of relating similarity and knowledge-based categorization principles.

Concepts as Organized by Theories

A number of researchers have argued that the organization of concepts is knowledge-based (rather than similarity-based) and driven by intuitive theories about the world (see, e.g., Carey, 1985; Gopnik & Meltzoff, 1997; Keil, 1986, 1989, 1995; Murphy & Medin, 1985; Rips & Collins, 1993; Schank, Collins, & Hunter, 1986; Waldmann, Holyoak, & Fratianne, 1995; Wattenmaker, 1995; see Komatsu, 1992; Murphy, 1993; Medin & Heit, 1999, for general reviews). Murphy and Medin suggested that the relation between a concept and an example is analogous to the relation between theory and data; that is, classification is not based simply on a direct matching of properties of the concept with those in the example, but rather requires that the example have the right "explanatory relationship" to the theory organizing the concept. Classification may be more like an inference process than like a similarity judgment. We may induce that a man is drunk because we see him jump into a pool fully clothed. If so, our determination is probably not because the property "jumps into pools while clothed" is directly listed with the concept *drunk*. Rather, it is because part of our concept of drunk involves a theory of impaired judgment that serves to explain the man's behavior.

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

case of a fire," the category becomes sensible (Barsalou, 1983). In addition, one could readily make judgments about whether new examples (e.g., personal papers, magazines) belonged to the category, judgments that would not be based on overall similarity to category members.

Susan Carey (1982, 1985) has shown that children's biological theories guide their conceptual development. In one study, 6-year-old children rated a toy monkey to be more similar to people than is a worm, but they also judged that the worm was more likely to have a spleen than was the toy monkey (a spleen was described as "a green thing inside people"). Although worms may be less similar to people than are toy monkeys, they are more similar in some respects, namely, common biological functions. And Carey's work shows that children's biological theories help them determine which respects are relevant. Thus, the 6-year-old children's rudimentary biological knowledge influences the structure of their concept of animal (see also Au & Romo, 1999; Coley, 1995; Gelman, 1996; Hatano & Inagaki, 1994; Inagaki, 1997; Keil, 1989; Simons & Keil, 1995, for other studies on the development of biological knowledge).

The idea that concepts might be knowledge-based rather than similarity-based suggests a natural way in which concepts may change—namely, through the addition of new knowledge and theoretical principles. We have a different set of categories for mental disorders now than we had 100 years ago, in part because our knowledge base has become more refined. Often knowledge of diseases develops from information about patterns of symptoms to a specification of underlying causes. For example, the advanced stages of syphilis were treated as a mental disorder until the causes and consequences of this venereal disease were better understood.

Putting Similarity in Its Place

How are theories and similarity related? Clearly, theories may affect similarity (Wisniewski & Medin, 1994). For example, people can now recognize that the fact that the planets revolve around the sun is similar to the fact that when something is dropped it falls down and not up, because both involve gravity. Interestingly, similarity may also act as a constraint on theories. The impact of perceptual similarity on the development of causal explanations is evident in the structure of people's everyday theories. Frazer's (1959) cross-cultural analysis of belief systems pointed to the widespread character of two principles; homeopathy and contagion. The principle of **homeopathy** is that causes and effects tend to be similar. One manifestation of this principle is homeopathic medicine, in which the cure (and the cause) are seen to resemble the symptoms. In the Azande culture, for example, the cure for ringworm is to apply fowl's excrement because the excrement looks like the ringworm. Shweder (1977) has provided strong support for the claim that resemblance is a fundamental conceptual tool of everyday thinking in all cultures, not just so-called primitive cultures. In our culture, people reject acceptable foods just because they are shaped into a form that represents a disgusting object (Rozin, Millman, & Nemeroff, 1986). We will spare you a concrete example.

Contagion is the principle that a cause must have some form of contact to transmit its effect. In general, the nearer two events are in time and space, the more likely they are to be perceived as causally related (see, e.g., Dickinson, Shanks, & Evenden, 1984; Michotte, 1963). Of course, even children recognize that the timing relations may be rather subtle. If children are exposed to a ball that hits a second ball and there is a delay before the second ball moves away, then they will not interpret the collision of the balls as causing the movement (Leslie, 1988), though the collision will be seen as causing the second ball to move if it starts to move right after the collision. People also tend to assume that causes and effects should be of similar magnitude. Einhorn and Hogarth (1986) pointed out that the germ theory of disease initially met with great resistance because people could not imagine how such tiny organisms could have such devastating effects.

It is important to recognize that homeopathy and contagion often point us in the right direction. Immunization can be seen as a form of homeopathic medicine that has an underlying theoretical principle supporting it. Our reading of these observations, however, is not that specific theoretical (causal) principles are constraining similarity but rather that similarity (homeopathy and contagion) acts as a constraint on the search for causal explanations. Even in classical conditioning studies, the similarity of the conditioned stimulus and the unconditioned stimulus can have a major influence on the rate of conditioning (Testa, 1974).

One way of integrating similarity and explanation is in terms of a notion of **psychological essentialism** (Gelman, Coley, & Gottfried, 1994; Keil, 1989; Medin & Ortony, 1989; Wattenmaker, Nakamura, & Medin, 1988). The main ideas are as follows: People act as if things (e.g., objects) have essences or underlying natures that make them the thing that they are. Essentialism seems to be an idea present in many cultures (e.g. Atran, 1990; Walker, 1992). For biological categories in our culture, people might identify essence with genetic structure. The essence constrains or generates (often external) properties that may vary in their centrality. For example, people in our culture believe that the categories male and female are genetically determined, but to pick someone out as male or female we rely on characteristics such as hair length, height, facial hair, and clothing that represent a mixture of secondary sexual characteristics and cultural conventions. Although these characteristics are less reliable than genetic evidence, they are far from arbitrary. Not only do they have some validity in a statistical sense but also they are tied to our biological and cultural conceptions of male and female.

Note that psychological essentialism refers not to how the world really is but rather to how people approach the world. Wastebaskets presumably have no true essence, although we may act as if they do. Both social and psychodiagnostic categories are at least partially culture-specific and, therefore, may represent constructions rather than discoveries about the world (see also Morey & McNamara, 1987).

Why should people act as if things had essences? Possibly the reason is that it may be a good strategy for learning about the world. Recall that categorization faces computational complexity problems and that organisms face a strong need to make correct and useful predictions and inferences on the basis of their categorization schemes. One could say that people adopt an essentialist strategy or heuristic,

namely, the generalization that things that look alike tend to share deeper properties (similarities) because it's typically an effective strategy. That is, our perceptual and conceptual systems appear to have evolved such that the essentialist heuristic is very often correct (Atran, 1990; Medin & Wattenmaker, 1987; Shepard, 1987). This is true even for human artifacts such as cars, computers, and camping stoves because structure and function tend to be correlated. Surface characteristics that are perceptually obvious or readily produced on feature-listing tasks may not so much constitute the core of a concept as point toward it. This observation suggests that classifying on the basis of similarity will be relatively effective much of the time, but that similarity will yield to knowledge of deeper principles. Thus, in the work of Gelman and Markman (1986) discussed earlier (Figure 10.8), category membership was more important than perceptual similarity in determining inductive inferences. Gelman and Wellman (1991) showed that even young children seem to use notions of essence in reasoning about biological kinds (see also Gelman & Hirschfeld, 1999). Susan Gelman has systematically traced the development of essentialism and its role in conceptual and linguistic development (see Gelman, 1998, for a review).

We have presented one way of relating similarity to knowledge structures and theories. Still, it would be misleading to state that there is any strong consensus on this general issue. Indeed, in the area of machine learning, a great deal of attention has been directed at the question of how to integrate similarity-based and explanation-based learning (see, e.g., Elman, 1989; Flann & Dietterich, 1989; Rajamoney & DeJong, 1987; Wisniewski & Medin, 1994). Note that research on theories shares with the work on perceptual learning the idea that similarity has an important role to play but that role is governed by the consequences of other processes that work to constrain similarity.

Are There Kinds of Categories?

Almost all of our discussion has been focused on object concepts and their structure. But what about abstract concepts such as *democracy* or *dilemma*, adjective concepts such as *daring*, or verb concepts such as *dance*? Do the generalizations concerning object concepts apply equally to these other kinds of concepts or are there different principles involved when one moves beyond object concepts? One motivation for an interest in kinds of categories is that a number of researchers, especially in the area of cognitive development have suggested that cognition is organized in terms of distinct domains, each characterized by (usually) innate constraints or skeletal principles of development. For example, naïve psychology (theories about people), naïve biology (theories about living things), and naïve physics (theories about the physical world) may constitute distinct domains with somewhat different principles of conceptual development (see Hirschfeld & Gelman, 1994, for examples).

It is transparently obvious that there are different kinds of categories, at least in the everyday sense of *different* and *kinds*. But to answer the sorts of questions we have raised, one must have some specific criteria in mind for what counts as a different kind. One might imagine that categories might differ in structure (as in our



A DEBATE

Category Specific Deficits

Studies of patients with brain damage often provide clues to normal brain functioning. One striking observation is that patients may show selective, **category-specific deficits** where they lose their ability to recognize and name category members from a particular domain of concepts. Nelson (1946) reported a patient who was unable to recognize a telephone, a hat, or a car but could identify people and other living things. Other researchers have noted the opposite deficit where people lose their ability to identify living kinds but retain the ability to categorize nonliving kinds.

Category-specific deficits seem consistent with ideas about domain-specific cognition but their interpretation has proven to be a matter of some controversy. On one side is the idea that living kinds and nonliving kinds are represented in anatomically and functionally distinct systems (Sartori & Job, 1988). An alternative view is that these deficit patterns can be accounted for by the fact that different kinds of information are needed to identify different kinds of objects (Warrington & Shallice, 1984). For example, sensory information may be relatively more important for recognizing living kinds and functional information more important for identifying nonliving kinds (e.g., artifacts). Although the evidence appear to be more consistent with the kinds of information view (see Damasio, Grabowski, Tranel, & Hichwa, 1996; Forde & Humpheys, 1999, for reviews), the issue continues to be hotly debated (see Caramazza & Shelton, 1998, for a vigorous defense of the domain specificity view of category specific deficits).

distinction between hierarchical and nonhierarchical categories), in the processing principles associated with them, or in principles that are tied to specific contents. The latter idea is tied to the notion of domain-specificity; the idea is that by looking only at principles that apply to all categories, we may be missing important principles that apply only to an important subset of categories (e.g., psychology vs. biology vs. physics).

Some of the candidate distinctions among kinds of categories, such as Barsalou's contrast between taxonomic and goal-derived categories have already been mentioned. We have space only to give a few examples of other contrasts. One distinction receiving increasing research interest is that between nouns and verbs, which may differ in structure (e.g., Macnamara, 1972; Gentner, 1981). One proposal is that nouns refer to clusters of correlated properties that create bounded chunks

of perceptual experience. Verbs generally focus on relations among these entities involving things such as causal relations, activity, or change of state. Since relations require objects, nouns may be conceptually simpler than verbs and more constrained by perceptual experience. If so, then we might expect more cross-linguistic variability in verbs than nouns and that syntactic structure should play a greater role in verb learning than in noun learning. There is evidence for both of these claims (e.g., Bowerman, 1996; Choi & Bowerman, 1991; Levinson, 1994; Naigles, 1990; Pinker, 1994). Be aware, however, that the distinction between nouns and verbs is more subtle than we have implied. For example, motion may be associated with both nouns and verbs (e.g., runner); nonetheless, there is a bias for nouns to be associated with motion intrinsic to an object and verbs to be associated with motions involving relations between objects (Kersten, 1998a, 1998b).

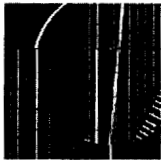
Even more attention has been directed at the idea of domain-specificity. Much of the work on essentialism has been conducted in the context of exploring children's naïve biology (see also Au, 1994; Carey, 1995; Gopnik & Wellman, 1994; Spelke, Phillips, & Woodward, 1995). Although it is difficult to give a precise definition of a domain, the notion of domain-specificity has served to organize a great deal of research on conceptual development. For example, there has been a strong focus on the question of whether and when young children distinguish between psychology and biology. Carey (1985) argues that young children understand biological concepts in terms of a naïve psychology where human beings are the prototypical psychological entity. Only later on do they reorganize their knowledge into a less humanocentric, biological form where human beings are simply one animal among many. Others (e.g., Keil, 1989) have argued that young children do have biologically specific theories, though not as elaborate as those of adults. The issue is not a matter of idle debate because understanding how children's biological knowledge develops is highly relevant to science education as well as children's health-related behaviors (e.g., Au & Romo, 1996).

Summary

Research on object concepts has received disproportionate attention in the psychology of concepts. This work suggests that people may use similarity and theories as organizing principles for their categories, though neither type of information is sufficient by itself to account for people's behavior with concepts. Indeed, in many cases, both similarity and theories are related, as theories may help determine what people consider similar, and physical similarity may influence the development of theories.

Other kinds of concepts (besides object concepts) may have different structure and different processing principles (e.g., mechanisms by which they are learned). In addition, it appears to be a useful research strategy to analyze conceptual development in terms of different domains. Overall, then, there is a trend to organize categorization principles in terms of specific kinds of concepts (see Medin, Lynch, & Solomon, 2000, for a general review).

USE OF CATEGORIES IN REASONING



So much attention has been paid to structural aspects of categories that only modest attention has been given to the question of how categories may be used in reasoning. The promising results from the few studies that have been done suggest that this question is worthy of much more attention.

Goals and Ad Hoc Categories

Earlier we mentioned Barsalou's work on goal-derived categories (1985, 1987, 1989). In particular he has studied the organization of categories constructed in the service of goals, which we will refer to as **ad hoc categories**. "Things to take on a camping trip" and "foods to eat while on a diet" are two such ad hoc categories. Barsalou has found that ad hoc categories show the same typicality or goodness-of-example effects that are seen with more established categories. As we noted before, however, the basis for these effects is not overall similarity of examples to each other or to a prototype, but rather similarity to an ideal. For example, typicality ratings for the category of things to eat on a diet is determined by how closely examples conform to the ideal of zero calories. If goal-derived categories are used repeatedly and consistently, they may become more stable in memory. In short Barsalou's work suggests that goals and the reasoning processes associated with them can affect category structure.

Conceptual Combination

As we mentioned at the beginning of this chapter, concepts provide "building blocks" or "mental tokens" that enable the construction of new concepts from old ones. For example, people use their knowledge of chocolate and rash to interpret *chocolate rash* as a rash caused by chocolate, even though they may never have seen the term chocolate rash before. **Conceptual combination** allows us to produce a virtually unlimited set of new concepts. Given that we hear novel combinations of concepts all the time, an important question is how people are able to understand them.

To our knowledge, a complete theory of how people combine concepts does not exist. Basically, we know that the most straightforward ideas are incomplete. One idea is that adjective-noun combinations are understood by constructing modified prototypes, the selective modification model of Smith & Osherson (1984). For example, according to their model, to understand the term *brown apple*, the apple prototype would be retrieved, the dimension of color would receive extra attention, and the prototypical value of red would be replaced with brown. In effect, one would be constructing a new prototype, brown apple, and the typicality of examples could be judged with respect to this constructed prototype. This modification model accounts for a number of phenomena associated with typicality judgments for combined adjective-noun concepts (Smith, Osherson, Rips, & Keane, 1988).

But conceptual combination is more complex than the modification model (or any other current model) implies. One problem is that the typicality of combined concepts cannot be predicted from the typicality of constituents. As an illustrative example, consider the concept *spoon*. People rate small spoons as more typical spoons than large spoons, and metal spoons as more typical spoons than wooden spoons. If the concept spoon is represented by a prototypic spoon, then a small metal spoon should be the most typical spoon, followed by small wooden and large metal spoons, and large wooden spoons should be the least typical. Instead, people find large wooden spoons to be more typical spoons than either small wooden spoons or large metal spoons (Medin & Shoben, 1988). One way for a prototype model to handle these results is to argue that people have separate prototypes for small spoons and large spoons. But this strategy creates new problems. Obviously one cannot have a separate prototype for every adjective-noun combination, because there are simply too many possible combinations. One might suggest that there are distinct subtypes for concepts like spoon, but one would need a theory describing how and when subtypes are created.

Another problem is that combined concepts may have properties that do not appear as properties of either constituent concept (Hampton, 1987, 1996; Murphy, 1988; see also Storms, De Boeck, Van Mechelen, & Ruts, 1998). For example, a salient property of the concept *pet bird* is that they live in cages, but this property is very atypical of either pets or birds by themselves.

A third problem is that people may complicate analyses by using a variety of strategies for comprehending combined concepts (Wisniewski, 1996, 1997; Wisniewski & Gentner, 1991). Consider, for example novel noun-noun combinations. When the two nouns are easily compared (because they have similar kinds of properties and are said to be "alignable") the combination is typically interpreted by extending a salient property of the modifying noun to the main noun. For example, a "zebra horse" might be understood as a horse with stripes. When the nouns are less comparable the most likely result is an integration where some relationship is found that meaningfully links the two concepts. For example, a "zebra house" typically is understood as a house for zebras rather than as a house with stripes. Models of conceptual combination that rely on a single process cannot account for these results (see Wisniewski & Love, 1998). So far, it seems that conceptual combination is as difficult as it is important (see Gagné & Shoben, 1997; Gerrig & Murphy, 1992; and Hampton, 1993, for other approaches to conceptual combination).

Categories and Induction

Finally, we note that categories play an important role in inductive reasoning. **Induction**, the use of some knowledge to draw inferences or expectations about other, more general situations, will be an important part of Chapter 11, but we include a sample here to illustrate the combined influence of similarity to examples and category membership in reasoning. Osherson, Smith, Wilkie, Lopez, and Shafir (1990) developed a formal model of category-based induction (see also Rips, 1975). The

main idea is that judgments are based on both similarity to examples and knowledge of category membership. The model is formulated in terms of "argument strength," going from a premise or premises to the conclusion (candidate inference). For example, a premise might be "robins have property x " and the conclusion "sparrows also have property x " or "all birds have property x ." We are using "have property x " to illustrate that properties were chosen so that participants in experiments would have no knowledge of whether these properties are true (e.g., "birds have sesame bones"). The task of the participants was to judge which argument form is stronger. For example, from the premise that robins have some property, people judge the argument with the conclusion that sparrows also have the property to be stronger than the argument with the conclusion that all birds have the property.

Argument strength is assumed to depend on two factors: the similarity of premise categories to the conclusion category and the similarity of the premise categories to examples of the lowest level category that the premise and conclusion categories share. For example, the strength of "robins have x " for the conclusion, "all birds have x ," is based on the similarity of robins to bird (assumed to be perfect, because robins are birds) and the similarity of robins to other birds. The net confidence is assumed to be based on some weighting of these two factors. Similarly, the strength of "ostriches have x " for the conclusion, "all birds have x ," is a function of the similarity of ostriches to bird (again, assumed to be perfect) and the similarity of ostriches to other birds. It follows directly that "robins have x " should be stronger than "ostriches have x " for the conclusion that all birds have x , because ostriches are less similar to other birds than are robins. In short, the model predicts typicality effects in reasoning (robins are more typical than ostriches).

Many of the predictions of the Osherson et al. model make intuitive sense, but a few predictions are quite surprising. The model predicts that going from the premise "bears have x " to the conclusion that "all mammals have x " will be stronger than going from "bears have x " to the conclusion "rabbits have x ," even though rabbits are mammals. Note that whatever is true of all mammals must also be true of rabbits. The reason for this prediction by the model is that going from bears to rabbits involves weighting separately the similarity of bears to rabbits (which may be low), whereas going from bears to mammals is based on the overall similarity of bears to other mammals (which should be fairly high because bears are typical mammals). The results agree with this counterintuitive prediction.

Another prediction of the model makes intuitive sense but the associated results depend very much on who the research participants are. The prediction concerns premise *diversity*. Consider the following arguments: "Suppose we discover two new diseases, A and B. So far we know that Disease A affects the white pine and the weeping willow and that Disease B affects the paper birch and the river birch. Which disease is more likely to affect all trees, A or B? The Osherson et al. model predicts that people will say A for the same reasons it predicts typicality effects. The two trees associated with Disease A are very different from each other and together they represent the category "tree" better (trees that the white pine isn't similar to the weeping willow may be similar to). You, the reader, probably would make the same judgment, perhaps because you think that Disease B might be specific to birches. Surprisingly,

however, when items like this (the items are different, because the plants and animals are different in Guatemala) are given to Itzá Maya adults, they do *not* pick the more diverse premises (Atran, 1998). And, in case you were starting to wonder about the reasoning processes of people in traditional societies, it also is the case that U.S. adults who know a lot about trees do not choose the disease associated with the more diverse pair consistently either (Coley, Medin, Proffitt, Lynch, & Atran, 1999). In fact, on the particular example given above tree maintenance workers overwhelmingly picked Disease B as more likely to affect all trees!

What accounts for these surprising results? The answer seems to be that both tree experts and Itzá Maya are reasoning about ecological relations and potential causal mechanisms. For example, a typical justification for choosing Disease B on the above item was that birches are very susceptible to disease and are very widely planted so there would be many opportunities for the disease to spread. Aside from suggesting that models of category-based reasoning need to make provision for people reasoning about potential mechanisms, these results point out that categorization and reasoning by novices in a domain may differ substantially from that of experts. If nothing else it should make cognitive psychologists more cautious about generalizing from results with undergraduate populations to all people.

Overall, the category-based induction model is quite successful. As we said at the beginning of this section, psychologists are just starting to explore the use of categories in reasoning (for extensions to nonblank properties, see Sloman, 1994; Smith, Shafir, & Osherson, 1993). The work on reasoning is part of a general trend to study how concepts are used and we can already see that this approach is delivering new insights into the nature of concepts.

Summary

Concepts serve to organize our mental life. Analyses of conceptual structure provide some important conclusions for our understanding of human cognition. Contrary to one's initial impression, many categories do not have defining properties but rather are fuzzy or probabilistic. These probabilistic categories may be organized around an average example (a prototype) or around specific examples seen during learning. The research we reviewed also suggests that things that are superficially similar tend to be similar in deeper ways, but similarity does not provide the basis for conceptual coherence. Concepts are often organized around theories and are knowledge-based rather than similarity-based.

Another important issue is that categories have many different functions. Much research on categorization has focused on how people learn to classify a set of items. However, people also use their categories to communicate, to reason, and to make predictions. Psychologists are now working on how category representations are used to serve these functions and also what people learn about categories in the process of using them.

What about the questions and problems with which we began this chapter? Why do we have the particular categories we have, out of the virtually unlimited

number of possibilities? Are categories in the world to be discovered, or are they imposed on the world by our minds? We have suggested that perceptual similarity may serve as an initial classification strategy that is refined, deepened, and itself modified by knowledge and developing theories about the world. If our perceptual and conceptual systems suggested to us categorization schemes that did not serve our goals and did not allow for predictions and explanations, we would be poorly adapted. Therefore, it is tempting to say that categories are organized around our goals and that different organisms with different goals would have very different categorization schemes. Although this may be true, it is equally important to recognize the importance of our environment. If we lived in a very different world, we would need a different perceptual and conceptual system to pick out categories that would be useful in that world. As indicated in our opening quotation, we believe that William James was correct when he suggested that mind and world “are something of a mutual fit.”

Key Terms

ad hoc categories	family resemblance
basic level	homeopathy
category	induction
category-specific deficits	linear separability
classical view	probabilistic view
concept	prototype
conceptual combination	psychological essentialism
contagion	typicality
exemplar models	

Recommended Readings

Because all three authors do research on categorization, we are tempted to say, “Read everything!” Fortunately, a number of reviews of the literature exist. For basic historical background, the edited volume by Rosch and Lloyd (1978) and the book by Smith and Medin (1981) are recommended. More recent analyses and overviews are provided by Estes (1994), Medin and Heit (1999), and the edited volumes by Hirschfeld and Gelman (1994), Lamberts and Shanks (1997), and Van Mechelen, Hampton, and Michalski (1993). A nice review of cross-cultural research on biological categorization is provided by Malt (1995) and the edited volume by Medin and Atran (1999) covers both cross-cultural and developmental research. Lillard (1998) argues that adult folk psychology differs across cultures—such a fact would have important implications for understanding the development of naïve or folk psychology. For analyses that focus on tests of or contrasts between particular models

of categorization, see Anderson (1990a, 1990b), Erickson and Kruschke (1998), Estes (1994), Heit (1994), Hintzman (1986), Maddox and Ashby (1993), Murphy and Ross (1994), Nosofsky et al. (1994), Nosofsky and Palmeri (1997), Storms, De Boeck, and Ruts (2000), and Wisniewski and Medin (1994). A special issue of the journal *Cognition*, edited by Sloman and Rips (1998), is devoted to the topic of the role of similarity in cognition. Malt, Sloman, Gennari, Shi, and Wang (1999) offer a nice cross-cultural analysis of the relation between perceived similarity of containers and how they are named.

Finally, there is increasing interest in the use of categories in reasoning. A good place to start is Rips (1975), Osherson et al. (1990), Smith et al. (1993), and Sloman (1993, 1994, 1998); for developmental work, see Gelman and Markman (1986), Gelman and Wellman (1991), Gutheil and Gelman (1997), Keil (1989), and Lopez, Gelman, Gutheil, and Smith (1992); for cross-cultural work see Atran (1998), Choi, Nisbett, and Smith (1997), and Lopez, Atran, Coley, Medin, and Smith (1997).

or are they
al similarity
d, and itself
r perceptual
id not serve
ld be poorly
l around our
different cat-
to recognize
ld, we would
s that would
believe that
re something

m

empted to say,
exist. For basic
) and the book
and overviews
ed volumes by
Van Mechelen,
arch on biolog-
e by Medin and
. Lillard (1998)
act would have
ve or folk psy-
rticular models