2

Typicality and the Classical View of Categories

The Classical View

When you were in junior high school or even earlier, you probably had a teacher who tried to get you to define words. Given a word like *wildebeest*, you might say "It's a kind of animal." Your teacher would not be satisfied with this definition, though. "A frog is also a kind of animal," she might point out, "Is it the same as a frog?" "No," you would reply sheepishly, "it has four legs and is about the size of a cow and has horns and lives in Africa." Well, now you're getting somewhere. My teacher used to tell us that a definition should include everything that is described by a word and nothing that isn't. So, if your description of a wildebeest picks out all wildebeest and nothing else, you have given a successful definition. (In fact, most definitions in dictionaries do not meet this criterion, but that is not our problem.)

Definitions have many advantages from a logical standpoint. For example, the truth or falsity of "Rachel is a wildebeest" is something that can be determined by referring to the definition: Does Rachel have all the properties listed in the definition—four legs, horns, and so on? By saying "Rachel is not a wildebeest," we are saying that Rachel lacks one or more of the definitional properties of wildebeest. And we could potentially verify statements like "Wildebeest can be found in Sudan" by looking for things in Sudan that meet the definition. Philosophers have long assumed that definitions are the appropriate way to characterize word meaning and category membership. Indeed, the view can be traced back as far as Aristotle (see Apostle 1980, pp. 6, 19–20). In trying to specify the nature of abstract concepts like fairness or truth, or even more mundane matters such as causality and biological kinds, philosophers have attempted to construct definitions of these terms. Once we have a definition that will tell us what exactly is and is not a cause, we will have come a long way toward understanding causality. And much philosophical argu-

Figure 2.1
Two concepts from Hull's (1920) study. Subjects learned 12 concepts like this at a time, with different characters intermixed. Subjects were to respond to each stimulus with the name at left (*oo* and *yer* in this case). Note that each example of a concept has the defining feature, the radical listed in the "concept" column.

mentation involves testing past definitions against new examples, to see if they work. If you can find something that seems to be a cause but doesn't fit the proposed definition, then that definition of cause can be rejected.

It is not surprising, then, that the early psychological approaches to concepts took a definitional approach. I am not going to provide an extensive review, but a reading of the most cited work on concepts written prior to 1970 reveals its assumption of definitions. I should emphasize that these writers did not always explicitly *say*, "I have a definitional theory of concepts." Rather, they took such an approach for granted and then went about making proposals for how people learned concepts (i.e., learned these definitions) from experience.

For example, Clark Hull's (1920) Ph.D. thesis was a study of human concept learning—perhaps surprisingly, given his enormous later influence as a researcher of simple learning, usually in rats. Hull used adapted Chinese characters as stimuli. Subjects viewed a character and then had to respond with one of twelve supposedly Chinese names (e.g., *oo* and *yer*). Each sign associated with a given name contained a radical or component that was identical in the different signs. As figure 2.1 shows, the *oo* characters all had the same radical: a kind of large check mark with two smaller marks inside it. Clearly, then, Hull assumed that every example of a concept had some element that was critical to it.

There are two aspects to a definition that these items illustrate. The first we can call *necessity*. The parts of the definition must be in the entity, or else it is not a member of the category. So, if a character did not have the check-mark radical, it would not be an *oo*. Similarly, if something doesn't have a distinctive attribute of chairs, it is not a chair. The second aspect we can call *sufficiency*. If something has all the parts mentioned in the definition, then it must be a member of the category. So, anything having that radical in Hull's experiment would be an *oo*, regardless of what other properties it had. Note that it is not enough to have one or two of the

parts mentioned in the definition—the item must have all of them: The parts of the definition are "jointly sufficient" to guarantee category membership. So, it wouldn't be enough for something to be as big as a cow and be from Africa to be a wildebeest; the item must also be an animal, with four legs and horns, if those things are included in our definition.

Hull (1920) explicitly adopted these aspects of a definition (without calling them by these names) in one of the principles he followed in creating his experiment (p. 13): "All of the individual experiences which require a given reaction, must contain certain characteristics which are at the same time common to all members of the group requiring this reaction [i.e., necessary] and which are NOT found in any members of the groups requiring different reactions [i.e., sufficient]." (Note that Hull adopted the behaviorist notion that it is the common response or "reaction" that makes things be in the same category. In contrast, I will be talking about the mental representations of categories.) Hull believed that this aspect of the experiment was matched by real-world categories. He describes a child who hears the word *dog* used in a number of different situations. "At length the time arrives when the child has a 'meaning' for the word dog. Upon examination this meaning is found to be actually a characteristic more or less common to all dogs and not common to cats, dolls and 'teddy-bears.' But to the child, the process of arriving at this meaning or concept has been largely unconscious. He has never said to himself 'Lo! I shall proceed to discover the characteristics common to all dogs but not enjoyed by cats and "teddy-bears"'" (p. 6). Note that although Hull says that "Upon *examination* this meaning is . . .," he does not describe any examination that has shown this, nor does he say what characteristics are common to all dogs. These omissions are significant—perhaps even ominous—as we shall soon see.

In the next major study of concept learning, Smoke (1932—the field moved somewhat more slowly in those days) criticized the definitional aspect of Hull's concepts. He says quite strongly that "if any concepts have ever been formed in such a fashion, they are very few in number. We confess our inability to think of a single one" (p. 3). He also quotes the passage given above about how children learn the concept of dog, and asks (pp. 3–4): "What, we should like to ask, is this 'characteristic more or less common to all dogs and not common to cats, dolls, and "teddy-bears"' that is the '"meaning" for the word dog' . . . What is the 'common element' in 'dog'? Is it something within the *visual* stimulus pattern? If exact drawings were made of all the dogs now living, or even of those with which any given child is familiar, would they 'contain certain strokes in common' [like Hull's characters] which could be 'easily observed imbedded in each'?"

One might think from this rather sarcastic attack on Hull's position that Smoke is going to take a very different position about what makes up a concept, one that is at odds with the definitional approach described. That, however, is not the case. Smoke's objection is to the notion that there is a single, simple element that is common to all category members. (Although Hull's stimuli do have such a simple element in common, it is not so clear that he intended to say that all categories were like this.) Smoke says "As one learns more and more about dogs, [one's] concept of 'dog' becomes increasingly rich, not a closer approximation to some bare 'element' " (p. 5). What Smoke feels is missing from Hull's view is that the essential components of a concept are a complex of features that are connected by a specified relationship, rather than being a single common element. Smoke gives this example of a concept called "zum," which he feels is more realistic: "Three straight red lines, two of which intersect the third, thereby trisecting it" (p. 9). You may judge for yourself whether this is more realistic than Hull's Chinese characters. In teaching such concepts, Smoke made up counterexamples that had some of the components of the concept but not all of them. Thus, for a nonzum, he might make up something with only two lines or something with three lines but that did not intersect one another in the required way. Thus, he created a situation that precisely follows the definitional view: Zums had all the required properties, and for the items that did not, subjects had to learn not to call them zums. Thus, the properties of zums were necessary and sufficient.

In short, although Smoke seems to be rejecting the idea of concepts as being formed by definitions, he in fact accepts this view. The main difference between his view and Hull's is that he viewed the definitions as being more complex than (he thought) Hull did. I have gone through these examples in part to show that one can tell whether experimenters had the definitional view by looking at their stimuli. In many older (pre-1970) studies of concepts, all the members have elements in common, which are not found in the nonmembers. Thus, even if such studies did not explicitly articulate this definitional view, they presupposed it.

The work of Hull and Smoke set the stage for research within American experimental psychology's study of concepts. In addition to their view of concepts, the techniques they developed for concept learning studies are still in use today. Another influence that promoted the use of definitions in the study of concepts was the work of Piaget in cognitive development. Piaget viewed thought as the acquisition of logical abilities, and therefore he viewed concepts as being logical entities that could be clearly defined. Again, Piaget did not so much argue for this view of concepts as simply assume it. For example, Inhelder and Piaget's (1964, p. 7) theory relied on

constructs such as: "the 'intension' of a class is the set of properties common to the members of that class, together with the set of differences, which distinguish them from another class"—that is, a definition. Following this, they provide a list of logical properties of categories which they argued that people must master in order to have proper concepts. (Perhaps not surprisingly, they felt that children did not have true concepts until well into school age—see chapter 10.) Piaget's work was not directly influential on most experimental researchers of adult concepts, but it helped to bolster the definitional view by its extraordinary influence in developmental psychology.

Main Claims of the Classical View

The pervasiveness of the idea of definitions was so great that Smith and Medin (1981) dubbed it the *classical view* of concepts. Here, then are the main claims of the classical view. First, concepts are mentally represented as definitions. A definition provides characteristics that are a) necessary and b) jointly sufficient for membership in the category. Second, the classical view argues that every object is either in or not in the category, with no in-between cases. This aspect of definition was an important part of the philosophical background of the classical view. According to the *law of the excluded middle*, a rule of logic, every statement is either true or false, so long as it is not ambiguous. Thus, "Rachel is a wildebeest" is either true or false. Of course, we may not know what the truth is (perhaps we don't know Rachel), but that is not the point. The point is that in reality Rachel either is or isn't a wildebeest, with no in-between, even if we don't know which possibility is correct. Third, the classical view does not make any distinction between category members. Anything that meets the definition is just as good a category member as anything else. (Aristotle emphasized this aspect of categories in particular.) An animal that has the feature common to all dogs is thereby a dog, just the same as any other thing that has that feature. In a real sense, the definition *is* the concept according to the classical view. So all things that meet the definition are perfectly good members of the concept, and all things that do not fit the definition are equally "bad" members (i.e., nonmembers) of the concept, because there is nothing besides the definition that could distinguish these things. As readers are no doubt thinking, this all sounds too good to be true.

Before confirming this suspicion, it is worthwhile to consider what kind of research should be done if the classical view is true. What is left to understand about the psychology of concepts? The basic assumption was that concepts were acquired by learning which characteristics were defining. However, unlike Hull's assumption (and more like Smoke's), it became apparent that any real-world concepts would

involve multiple features that were related in a complex way. For example, dogs have four legs, bark, have fur, eat meat, sleep, and so on. Some subset of these features might be part of the definition, rather than only one characteristic. Furthermore, for some concepts, the features could be related by rules, as in the following (incomplete) definition of a strike in baseball: "the ball must be swung at and missed OR it must pass above the knees and below the armpits and over home plate without being hit OR the ball must be hit foul (IF there are not two strikes)." The use of logical connectives like AND, OR, and IF allows very complex concepts to be defined, and concepts using such connectives were the ones usually studied in psychology experiments. (There are not that many experiments to be done on how people learn concepts that are defined by a single feature.) Often these definitions were described as *rules* to tell what is in a category.

Bruner, Goodnow, and Austin (1956) began the study of this sort of logically specified concept, and a cottage industry sprang up to see how people learned them. In such experiments, subjects would go through many cards, each of which had a picture or description on it. They would have to guess whether that item was in the category or not, usually receiving feedback on their answer. For difficult concepts, subjects might have to go through the cards a dozen times or more before reaching perfect accuracy on categorizing items; for easy concepts, it might only take four or five run-throughs (blocks). Research topics included how learning depended on the number of attributes in the rule, the relations used in the rule (e.g., AND vs. OR), the way that category examples were presented (e.g., were they shown to the subject in a set order, or could the subject select the ones he or she wished?), the complexity of the stimuli, and so on. But whether this research is of interest to us depends on whether we accept the classical view. If we do not, then these experiments, although valid as examples of learning certain abstract rules, do not tell us about how people learn normal concepts. With that foreshadowing, let us turn to the problems raised for the classical view.

Problems for the Classical View

The groundbreaking work of Eleanor Rosch in the 1970s essentially killed the classical view, so that it is not now the theory of any actual researcher in this area (though we will see that a few theorists cling to it still). That is a pretty far fall for a theory that had been the dominant one since Aristotle. How did it happen? In part it happened by virtue of theoretical arguments and in part by the discovery of data that could not be easily explained by the classical view. Let us start with the argu-

ments. Because this downfall of the classical view has been very ably described by Smith and Medin (1981) and others (e.g., Mervis and Rosch 1981; Rosch 1978), I will spend somewhat less time on this issue than could easily be spent on it. Interested readers should consult those reviews or the original articles cited below for more detailed discussion.

In-Principle Arguments

The philosopher Wittgenstein (1953) questioned the assumption that important concepts could be defined. He used the concept of a game as his example. It is maddeningly difficult to come up with a definition of games that includes them but does not include nongame sports (like hunting) or activities like kicking a ball against a wall. Wittgenstein urged his readers not to simply say "There *must* be something in common," but to try to specify the things in common. Indeed, it turns out to be very difficult to specify the necessary and sufficient features of *most* real-world categories. So, the definition we gave earlier of dogs, namely, things that have four legs, bark, have fur, eat meat, and sleep, is obviously not true. Does something have to have four legs to be a dog? Indeed, there are unfortunate dogs who have lost a leg or two. How about having fur? Although most dogs do have fur, there are hairless varieties like chihuahuas that don't. What about barking? Almost all dogs bark, but I have in fact known a dog that "lost" its bark as it got older. This kind of argument can go on for some time when trying to arrive at necessary features. One can find some features that seem a bit "more necessary" than these—abstract properties such as being animate and breathing appear to be true of all dogs (we are talking about real dogs here, and not toy dogs and the like). However, although these features may be necessary, they are not nearly sufficient. In fact, all animals are animate and breathe, not just dogs. So, features like these will not form adequate definitions for dogs—they won't separate dogs from other similar animals. Also, sometimes people will propose features such as "canine" as being definitional of dogs. However, this is just cheating. "Canine" is simply a synonym for the dog concept itself, and saying that this is part of the definition of dog is not very different from saying that the dogs all have the property of "dogginess." Clearly, such circular definitions explain nothing.

Wittgenstein's argument, which is now widely accepted in the field, does have a problem, however (Smith and Medin 1981): It is primarily negative. It says something like "I can't think of any defining features for games or dogs," but this does not prove that there aren't any. Perhaps it's just that we are not clever enough to think of the defining features. Perhaps there are defining features that will be

realized or discovered in 100 years. This may be true, but it is incumbent on someone who believes the classical view to explain what the defining features are, *and* why we can't easily think of them. If our concept of dogs is a definition, why are we so bad at saying what it is even when we know the concept? Why is it that we can use this definition for identifying dogs and for thinking about them, but the properties we give for dogs are not definitional? The classical view has considerable trouble explaining this.

I used to think that the arguments against the classical view might be somewhat limited. In particular, it seemed likely to me that in certain technical domains, concepts might be well defined. For example, the rules of baseball are written down in black and white, and they look very much like the rules of old category-learning experiments (think of the disjunctive rule for a strike). Perhaps in the physical sciences, one will find classical concepts, as the scientists will have figured out the exact rules by which to decide something is a metal or a member of a biological genus. However, my own experience has always been that whenever one explores these domains in greater depth, one finds more and more fuzziness, rather than perfectly clear rules.

For example, consider the following portion of a lecture on metals by a distinguished metallurgist (Pond 1987). He begins by attempting, unsuccessfully, to get audience members to define what a metal is.

> Well, I'll tell you something. You really don't know what a metal is. And there's a big group of people that don't know what a metal is. Do you know what we call them? Metallurgists! . . . Here's why metallurgists don't know what metal is. We know that a metal is an element that has metallic properties. So we start to enumerate all these properties: electrical conductivity, thermal conductivity, ductility, malleability, strength, high density. Then you say, how many of these properties does an element have to have to classify as a metal? And do you know what? We can't get the metallurgists to agree. Some say three properties; some say five properties, six properties. We really don't know. So we just proceed along presuming that we are all talking about the same thing. (pp. 62–63)

And in biology, biologists are constantly fighting about whether things are two different species or not, and what species belong in the same genus, and so on. There is no defining feature that identifies biological kinds (Mayr 1982). In 2000, there was a dispute over whether Pluto is really a planet (apparently, it isn't, though it is too late to do anything about it now). Students in the town of Pluto's discoverer put up a web site demanding Pluto's retention as a full-fledged planet. Among their points was that the astronomers critical of Pluto's planethood "don't even have a definition of what a planet really is!" (Unfortunately, they did not provide their own definition.) The very fact that astronomers can argue about such cases is evidence that

the notion of planet is not well defined. The idea that all science consists of hard-and-fast logical categories, in contrast to those of everyday life, may be a romantic illusion.

One might well hope that legal concepts have a classical nature, so that definitions can be evenly applied across the board. One does not want judges and juries having to make decisions in fuzzy cases where there is no clear boundary between the legal and illegal. However, legal practice has found that in practice the law is sometimes very fuzzy, and legal theorists (e.g., Hart 1961) suggest that this is inevitably the case, because lawmakers cannot foresee all the possibilities the law will have to address, some of which do not exist at the time the law is passed (e.g., laws on intellectual property were written before the invention of the internet). Furthermore, since laws are written in language that uses everyday concepts, to the degree that these concepts are fuzzy, the law itself must be fuzzy. So, if one has a rule, "No vehicles in the park" (Hart 1961), does one interpret that to include wheelchairs? Maintenance vehicles? Ambulances? Bicycles? Although the law is stated in a simple, clear-cut manner, the fuzziness of the vehicle category prevents it from being entirely well-defined.[1]

Even artificial domains may have rules that are not particularly well defined. In 1999, Major League Baseball made a much publicized effort to clean up and standardize the strike zone, suggesting that my belief that strikes were well defined prior to that was a delusion. Perhaps the only hope for true classical concepts is within small, closed systems that simply do not permit exceptions and variation of the sort that is found in the natural world. For example, the rules of chess may create classical concepts, like bishops and castling. Baseball, on the other hand, has too much human behavior and judgment involved, and so its categories begin to resemble natural categories in being less well defined than the purist would like.

Empirical Problems

Unfortunately for the classical view, its empirical problems are even greater than its theoretical ones. One general problem is that the neatness envisioned by the classical view does not seem to be a characteristic of human concepts. As mentioned above, the notion of a definition implies that category membership can be discretely determined: The definition will pick out all the category members and none of the nonmembers. Furthermore, there is no need to make further distinctions among the members or among the nonmembers. In real life, however, there are many things that are not clearly in or out of a category. For example, many people express uncertainty about whether a tomato is a vegetable or a fruit. People are not sure about

whether a low, three-legged seat with a little back is a chair or a stool. People do not always agree on whether sandals are a kind of shoe. This uncertainty gets even worse when more contentious categories in domains such as personality or aesthetics are considered. Is *Sergeant Pepper's Lonely Hearts Club Band* a work of art? Is your neighbor just shy or stuck up? These kinds of categorizations are often problematic.

Research has gone beyond this kind of anecdote. For example, Hampton (1979) asked subjects to rate a number of items on whether they were category members for different categories. He did not find that items were segregated into clear members and nonmembers. Instead, he found a number of items that were just barely considered category members and others that were just barely not members. His subjects just barely included sinks as members of the kitchen utensil category and just barely excluded sponges; they just included seaweed as a vegetable and just barely excluded tomatoes and gourds. Indeed, he found that for seven of the eight categories he investigated, members and nonmembers formed a continuum, with no obvious break in people's membership judgments.

Such results could be due to disagreements among different subjects. Perhaps 55% of the subjects thought that sinks were clearly kitchen utensils and 45% thought they were clearly not. This would produce a result in which sinks appeared to be a borderline case, even though every individual subject had a clear idea of whether they were category members or not. Thus, the classical view might be true for each individual, even though the group results do not show this. However, McCloskey and Glucksberg (1978) were able to make an even stronger argument for such unclear cases. They found that when people were asked to make repeated category judgments such as "Is an olive a fruit?" or "Is a dog an animal?" there was a subset of items that individual subjects changed their minds about. That is, if you said that an olive was a fruit on one day, two weeks later you might give the opposite answer. Naturally, subjects did not do this for cases like "Is a dog an animal?" or "Is a rose an animal?" But they did change their minds on borderline cases, such as olive-fruit, and curtains-furniture. In fact, for items that were intermediate between clear members and clear nonmembers, McCloskey and Glucksberg's subjects changed their mind 22% of the time. This may be compared to inconsistent decisions of under 3% for the best examples and clear nonmembers (see further discussion below). Thus, the changes in subjects' decisions do not reflect an overall inconsistency or lack of attention, but a bona fide uncertainty about the borderline members. In short, many concepts are not clear-cut. There are some items that one cannot make up one's mind about or that seem to be "kind of" members. An avo-

The Necessity of Category Fuzziness

The existence of unclear examples can be understood in part as arising from the great variation of things in the world combined with the limitations on our concepts. We do not wish to have a concept for every single object—such concepts would be of little use and would require enormous memory space. Instead, we want to have a fairly small number of concepts that are still informative enough to be useful (Rosch 1978). The ideal situation would probably be one in which these concepts did pick out objects in a classical-like way. Unfortunately, the world is not arranged so as to conform to our needs.

For example, it may be useful to distinguish chairs from stools, due to their differences in size and comfort. For the most part, we can distinguish the two based on the number of their legs, presence of a back or arms, and size. However, there is nothing to stop manufacturers from making things that are very large, comfortable stools; things that are just like chairs, only with three legs; or stools with a back. These intermediate items are the things that cause trouble for us, because they partake of the properties of both. We could try to solve this by forming different categories for stools with four legs (stegs), for chairs with three legs (chools), stools with backs (stoocks), stools with backs and arms (stoorms), and so on. But by doing so, we would end up increasing our necessary vocabulary by a factor of 5 or 10, depending on how many distinctions we added for every category. And there would still be the problem of intermediate items, as manufacturers would no doubt someday invent a combination that was between a stoock and a stoorm, and that would then be an atypical example of both. Just to be difficult, they would probably also make stools with no back, with very tiny backs, with somewhat tiny backs, . . . up to stools with enormous, high backs. Thus, there would be items in between the typical stools and stoocks where categorization would be uncertain.

The gradation of properties in the world means that our smallish number of categories will never map perfectly onto all objects: The distinction between members and nonmembers will always be difficult to draw or will even be arbitrary in some cases. If the world existed as much more distinct clumps of objects, then perhaps our concepts could be formed as the classical view says they are. But if the world consists of shadings and gradations and of a rich mixture of different kinds of properties, then a limited number of concepts would almost have to be fuzzy.

cado is "kind of a vegetable," even if it is not wholeheartedly a vegetable. The classical view has difficulty explaining this state of affairs; certainly, it did not predict it.

Another problem for the classical view has been the number of demonstrations of *typicality effects*. These can be illustrated by the following intuition. Think of a fish, any fish. Did you think of something like a trout or a shark, or did you think of an eel or a flounder? Most people would admit to thinking of something like the first: a torpedo-shaped object with small fins, bilaterally symmetrical, which swims in the water by moving its tail from side to side. Eels are much longer, and they slither; flounders are also differently shaped, aren't symmetrical, and move by waving their body in the vertical dimension. Although all of these things are technically fish, they do not all seem to be equally good examples of fish. The *typical* category members are the good examples—what you normally think of when you think of the category. The *atypical* objects are ones that are known to be members but that are unusual in some way. (Note that *atypical* means "not typical," rather than "a typical example." The stress is on the first syllable.) The classical view does not have any way of distinguishing typical and atypical category members. Since all the items in the category have met the definition's criteria, all are category members. (Later I will try to expand the classical view to handle this problem.)

What is the evidence for typicality differences? Typicality differences are probably the strongest and most reliable effects in the categorization literature. The simplest way to demonstrate this phenomenon is simply to ask people to rate items on how typical they think each item is of a category. So, you could give people a list of fish and ask them to rate how typical each one is of the category fish. Rosch (1975) did this task for 10 categories and looked to see how much subjects agreed with one another. She discovered that the reliability of typicality ratings was an extremely high .97 (where 1.0 would be perfect agreement)—though later studies have suggested that this is an overestimate (Barsalou 1987). In short, people agree that a trout is a typical fish and an eel is an atypical one.

But does typicality affect people's use of the categories? Perhaps the differences in ratings are just subjects' attempt to answer the question that experimenters are asking. Yes, a trout is a typical fish, but perhaps this does not mean that trouts are any better than eels in any other respect. Contrary to this suggestion, typicality differences influence many different behaviors and judgments involving concepts. For example, recall that I said earlier that McCloskey and Glucksberg (1978) found that subjects made inconsistent judgments for only a subset of their items. These items could be predicted on the basis of typicality. Subjects did not change their minds about the very typical items or the clear nonitems, but about items in the middle of the scale, the atypical members and the "close misses" among the nonmembers. For

example, waste baskets are rated as atypical examples of furniture (4.70, where 1 is low and 10 is high), and subjects changed their minds about this item a surprising 30% of the time. They never changed their minds about tables, a very typical member (rated 9.83), or windows, a clear nonmember (rated 2.53).

Rips, Shoben, and Smith (1973) found that the ease with which people judged category membership depended on typicality. For example, people find it very easy to affirm that a robin is a bird but are much slower to affirm that a chicken (a less typical item) is a bird. This finding has also been found with visual stimuli: Identifying a picture of a chicken as a bird takes longer than identifying a pictured robin (Murphy and Brownell 1985; Smith, Balzano, and Walker 1978). The influence of typicality is not just in identifying items as category members—it also occurs with the production of items from a category. Battig and Montague (1969) performed a very large norming study in which subjects were given category names, like *furniture* or *precious stone* and had to produce examples of these categories. These data are still used today in choosing stimuli for experiments (though they are limited, as a number of common categories were not included). Mervis, Catlin and Rosch (1976) showed that the items that were most often produced in response to the category names were the ones rated as typical (by other subjects). In fact, the average correlation of typicality and production frequency across categories was .63, which is quite high given all the other variables that affect production.

When people learn artificial categories, they tend to learn the typical items before the atypical ones (Rosch, Simpson, and Miller 1976). Furthermore, learning is faster if subjects are taught on mostly typical items than if they are taught on atypical items (Mervis and Pani 1980; Posner and Keele 1968). Thus, typicality is not just a feeling that people have about some items ("trout good; eels bad")—it is important to the initial learning of the category in a number of respects. As we shall see when we discuss the explanations of typicality structure, there is a very good reason for typicality to have these influences on learning.

Learning is not the end of the influence, however. Typical items are more useful for inferences about category members. For example, imagine that you heard that eagles had caught some disease. How likely do you think it would be to spread to other birds? Now suppose that it turned out to be larks or robins who caught the disease. Rips (1975) found that people were more likely to infer that other birds would catch the disease when a typical bird, like robins, had it than when an atypical one, like eagles, had it (see also Osherson et al. 1990; and chapter 8).

As I will discuss in chapter 11, there are many influences of typicality on language learning and use. Just to mention some of them, there are effects on the order of word production in sentences and in comprehension of anaphors. Kelly, Bock, and

Keil (1986) showed that when subjects mentioned two category members together in a sentence, the more typical one is most likely to be mentioned first. That is, people are more likely to talk about "apples and limes" than about "limes and apples." Garrod and Sanford (1977) asked subjects to read stories with category members in them. For example, they might read about a goose. Later, a sentence would refer to that item with a category name, such as "The bird came in through the front door." Readers took longer to read this sentence when it was about a goose (an atypical bird) than when it was about a robin (a typical bird). Rosch (1975) found that typical items were more likely to serve as *cognitive reference points*. For example, people were more likely to say that a patch of off-red color (atypical) was "virtually" the same as a pure red color (typical) than they were to say the reverse. Using these kinds of nonlinguistic stimuli showed that the benefit of the more typical color was not due to word frequency or other aspects of the words themselves. Similarly, people prefer to say that 101 is virtually 100 rather than 100 is virtually 101.

This list could go on for some time. As a general observation, one can say that whenever a task requires someone to relate an item to a category, the item's typicality influences performance. This kind of result is extremely robust. In fact, if one compares different category members and does *not* find an effect of typicality, it suggests that there is something wrong with—or at least unusual about—the experiment. It is unfortunate for the classical view, therefore, that it does not predict the most prevalent result in the field. Even if it is not specifically disproved by typicality effects (see below), it is a great shortcoming that the view does not actually explain why and how they come about, since these effects are ubiquitous.

Revision of the Classical View

As a result of the theoretical arguments and the considerable evidence against the classical view, a number of writers have attempted to revise it so that it can handle the typicality data and unclear members. The main way to handle this has been to make a distinction between two aspects of category representation, which I will call the *core* and *identification procedures* (following Miller and Johnson-Laird 1976; see Armstrong, Gleitman, and Gleitman 1983; Osherson and Smith 1981; Smith and Medin 1981; and Smith, Rips, and Shoben 1974 for similar ideas). The basic idea is as follows. Although concepts do have definitions (which we have not yet been able to discover), people have also learned other things about them that aren't definitional. This kind of information helps us to identify category members or to

use information that is not defining. For example, not all dogs have fur, so having fur cannot be part of the definition of the dog category. However, it is still useful to use fur as a way of identifying dogs, because so many of them do have it. Thus, "fur" would be part of the identification procedure by which we tell what actual dogs are, but it would not be part of the concept core, which contains only the definition. One could call "fur" a *characteristic feature*, since it is generally true of dogs even if not always true: Characteristically, dogs have fur.

Part of the problem with this proposal is that it is not clear what the concept core is supposed to do. If it isn't used to identify the category members, then what is it for? All the typicality effects listed above must arise from the identification procedure, since the category core by definition (sic) does not distinguish typicality of members. One proposal is that people use the identification procedure for fast and less reliable categorization, but that they will use the category core for more careful categorization or reasoning (e.g., Armstrong et al. 1983; Smith et al. 1974). Thus, when tasks involve more careful thought, with less time pressure than in many experiments, people might be encouraged to use the category core more. For example, Armstrong et al. (1983) found that people took longer to identify less typical even numbers than more typical ones (e.g., 4 is a more typical even number than 38 is). However, since subjects know the rule involving even numbers and are extremely accurate at this, they may use the category core to ultimately decide the question. Whether this argument can be extended to items that do not have such a clear rule, of course, needs to be considered.

In summary, on this revised view, the effects of typicality result from the identification procedures, whereas certain other behaviors (primarily categorization decisions) depend primarily on the concept core.

I will criticize this theory at the end of the chapter. However, there have been some empirical tests of this proposal as well. First, Hampton's (1979) study mentioned above also included a component in which subjects listed properties of different categories, and he attempted to separate defining from characteristic features. For example, subjects first said what properties they used to decide that something was a fruit. Other subjects then evaluated examples of fruit and similar objects to see which ones had the properties. So, they would have considered whether apple and avocado have properties such as "is from a plant," "is eaten," and "has an outer layer of skin or peel," which were mentioned by other subjects as being critical features. Hampton derived a list of necessary features for each category, by including the listed features that were found in all the category members. For example, all the items that his subjects identified as fruit had the feature "is eaten," and so this

was a necessary feature. The question he next asked was whether these features were defining: If an item had all these features, would it be in the category? The answer is no. He found many cases of items that had all of these necessary features but were not considered to be category members. For example, cucumbers and garlic had all of the necessary features for fruit but were not considered to be category members. This, then, is another failure to find defining features of actual categories. Furthermore, Hampton found that when he simply counted up how many relevant features each item had (not just the necessary features, but all of them), he could predict how likely people were to include the item as a category member. But since all members would be expected to have the defining features (according to the revised classical view), the number of other features should not predict category membership. Thus, nondefining features are important in deciding category membership—not just core features.

In more recent work, Hampton (1988b, 1995) has examined the relationship between typicality measures and category membership judgments. According to the revised classical view, typicality is not truly related to category membership but simply reflects identification procedures. So extremely atypical items like whales or penguins are just as much mammals and birds, respectively, as typical examples are (Armstrong et al. 1983; Osherson and Smith 1981). However, Hampton's results in a number of domains show that typicality ratings are the best predictor of untimed category judgments, the ones that should only involve the core. These results appear to directly contradict the revised classical view.

One reason for the popularity of the classical view has been its ties to traditional logic (Inhelder and Piaget 1964). For example, how does one evaluate sentences of the sort "All dogs are animals" or "Coach is a dog and a pet"? Propositional logic tells us that "Coach is a dog and a pet" is true if "Coach is a dog" and "Coach is a pet" are both true. This can be easily accommodated in the classical view by the argument that "Coach is a dog and a pet" is true if Coach has the necessary and sufficient features of both dogs and pets. Surprisingly, there is empirical evidence suggesting that people do not follow this rule. Hampton (1988b) found that people are willing to call something a member of a conjunctive category (X AND Y) even if it is not in both components (X, Y). For example, subjects believe that chess is in the category sports that are also games, but they do not believe that it is a sport. So, chess seems to fulfill the definition for sport in one context but not in another. He also found (Hampton 1997) that subjects believed that tree houses are in the category of dwellings that are not buildings, but they also believe them to be buildings. So, on different occasions, people say that it does and does not have the defining

features of buildings. Although very troublesome for the classical view, these examples have a very natural explanation on other views, as is explained in chapter 12.

A related advantage that has been proposed for the classical view is that it has a very natural way of explaining how categories can be *hierarchically ordered*. By this I mean that categories can form nested sets in which each category includes the ones "below" it. For example, a single object could be called Coach (his name), a yellow labrador retriever, a labrador retriever, a retriever, a dog, a mammal, a vertebrate, and an animal. These particular categories are significant because *all* yellow labs are dogs, all dogs are mammals, all vertebrates are animals, and so on. As we will see in chapter 7, this aspect of categories has been thought to be quite significant. The classical view points out that if all *X* are *Y*, then the definition of *Y* must be included in the definition of *X*. For example, all red triangles are triangles. Therefore, red triangles must be closed, three-sided figures, because this is the definition of a triangle. Similarly, whatever the definition of labradors is, that must be included in the definition of yellow labs, because all yellow labs are labradors. This rule ensures that category membership is *transitive*. If all *A*s are *B*s, and all *B*s are *C*s, then all *A*s must be *C*s. Since the definition of *C* must be included in *B* (because all *B*s are *C*s), and the definition of *B* must be included in *A* (because all *A*s are *B*s), the definition of *C* must thereby be included in *A*. The nesting of definitions provides a way of explaining how categories form hierarchies.

Hampton (1982) suspected that there might be failures of transitivity, which would pose a significant problem for the classical view. He asked subjects to decide whether items were members of two categories—one of them a subset of the other. For example, subjects decided whether an armchair is a chair and whether it is furniture. They also had to judge whether chairs are furniture (they are). Hampton found a number of cases in which an item was judged to be a member of the subset category but not the higher category—that is, examples of chairs that were not thought to be furniture. For instance, subjects judged that chairs were furniture and that a car seat was a chair; however, they usually denied that a car seat was furniture. But if a car seat has the defining features of chairs, and chairs have the defining features of furniture, then car seat must have the defining features of furniture. It should be pointed out that Hampton's task was not a speeded, perceptual judgment, but a more leisurely decision about category membership, which is just the sort of judgment that should involve the concept core. It is puzzling to the revised classical view that even such judgments do not show the use of definitions in the way that is expected. However, we will see later that this kind of intransitivity is easily explained by other views.

Finally, a theoretical problem with the revised classical view is that the concept core does not in general appear to be an important part of the concept, in spite of its name and theoretical intention as representing the "real" concept. As mentioned earlier, almost every conceptual task has shown that there are unclear examples and variation in typicality of category members. Because the concept core does not allow such variation, all these tasks must be explained primarily by reference to the identification procedure and characteristic features. So, if it takes longer to verify that a chicken is a bird than that a bluejay is a bird, this cannot be explained by the concept core, since chicken and bluejay equally possess the core properties of birds, according to this view. Instead, chicken and bluejays differ in characteristic features, such as their size and ability to fly. Thus, speeded judgments must not be relying on the category core. When this reasoning is applied to all the tasks that show such typicality effects, including category learning, speeded and unspeeded categorization, rating tasks, language production and comprehension, vocabulary learning, and category-based induction, the concept core is simply not explaining most of the data. As a result, most researchers have argued that the concept core can simply be done away with, without any loss in the ability to explain the results (see especially Hampton 1979, 1982, 1995).

Summary of Typicality as a Phenomenon

Before going on to the more recent theoretical accounts of typicality phenomena, it is useful to discuss these phenomena in a theory-neutral way. Somewhat confusingly, the phenomena are often referred to as revealing a *prototype structure* to concepts. (This is confusing, because there is a prototype *theory* that is a particular theory of these results, so sometimes *prototype* refers to the phenomenon and sometimes the specific theory. This is not an isolated case of terminological confusion in the field, as you will see.) A prototype is the best example of a category. One can think of category members, then, arranged in order of "goodness," in which the things that are very similar to the prototype are thought of as being very typical or good members, and things that are not very similar as being less typical or good members (Rosch 1975).

One way to illustrate this concept concretely is by the dot-pattern studies of Posner and Keele (1968, 1970), since it is very clear what the prototype is in their experiments. (Also, these are historically important experiments that are of interest in their own right.) Posner and Keele first generated a pattern of randomly positioned dots (see figure 2.2a) as the category prototype. From this pattern, they made

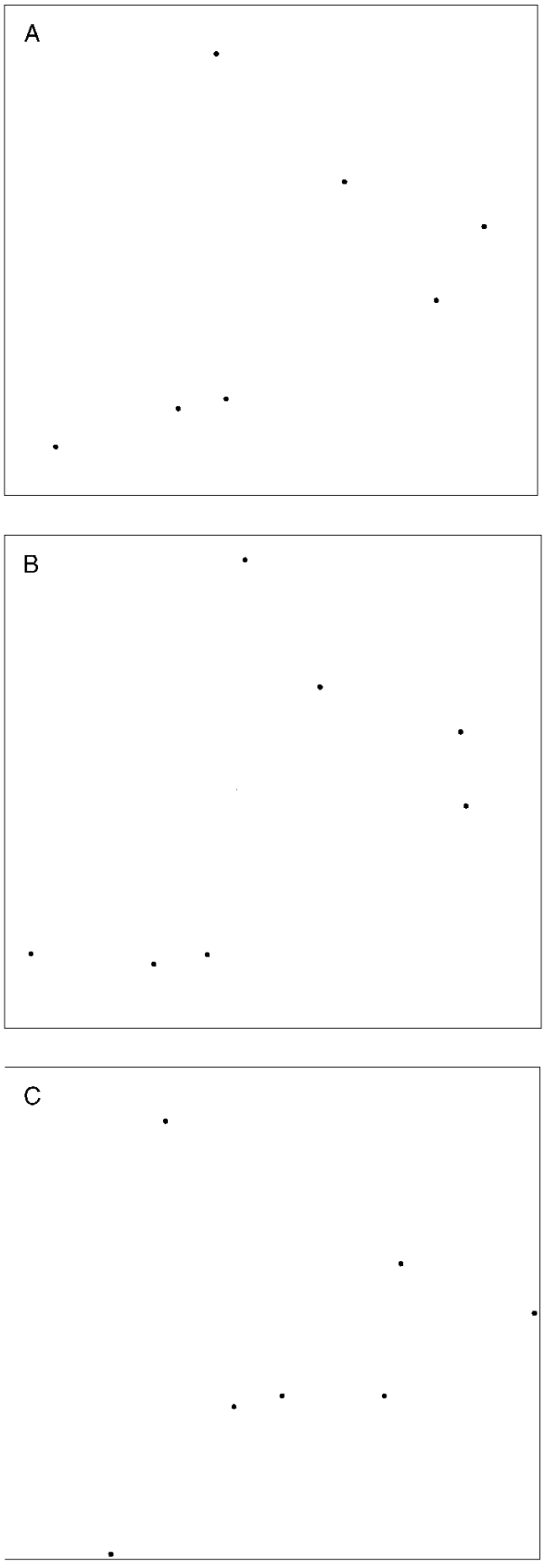

Figure 2.2
Dot patterns of the sort used by Posner and Keele (1968, 1970). 2.2a represents the initial randomly placed dots, the prototype. 2.2b represents a small distortion of the prototype, in which each dot is moved a small distance in a random direction. 2.2c represents a large distortion, in which each dot is moved a large distance. On can make indefinitely many such stimuli by generating a different (random) direction for each dot.

many more patterns of dots, by moving each point in the original pattern in a random direction. In some cases, they moved the points a small amount in that random direction (as in figure 2.2b), and for other items, they moved the points a large amount (as in figure 2.2c). These new patterns were called "distortions" of the original. Here, it is clear that the prototype is the most typical, best member of the category, because it was the basis for making the other patterns. However, it is also the case that the distortions were all somewhat similar to one another by virtue of the fact that they were constructed from the same prototype. Indeed, subjects who viewed only the distortions of four prototypes were able to learn the categories by noticing this similarity (they had no idea how the patterns were made). Furthermore, the prototype itself was identified as a member of the category in test trials, even though subjects had never seen it during learning.[2] Finally, the items that were made from small distortions were learned better than items made from large distortions. And when subjects were tested on new items that they had not specifically been trained on, they were more accurate on the smaller distortions. Figure 2.2 illustrates this nicely: If you had learned that a category looked generally like the pattern in 2.2a (the prototype), you would find it easier to classify 2.2b into that category than you would 2.2c. In summary, the smaller the distortion, the more typical an item was, and the more accurately it was classified.

This example illustrates in a very concrete way how prototypes might be thought of in natural categories. Each category might have a most typical item—not necessarily one that was specifically learned, but perhaps an average or ideal example that people extract from seeing real examples. You might have an idea of the prototypical dog, for example, that is average-sized, dark in color, has medium-length fur, has a pointed nose and floppy ears, is a family pet, barks at strangers, drools unpleasantly, and has other common features of dogs. Yet, this prototype may be something that you have never specifically seen—it is just your abstraction of what dogs are most often like. Other dogs vary in their similarity to this prototype, and so they differ in their typicality. Miniature hairless dogs are not very similar to the prototype, and they are considered atypical; St. Bernards are not very similar, albeit in a different way, and so they are also atypical; a collie is fairly similar to the prototype, so it is considered fairly typical; and so on. As items become less and less similar to the prototype, it is more and more likely that they won't be included in the category. So, if you saw a thing with four legs but a very elongated body, no hair, and whiskers, you might think that it was somewhat similar to a dog, but not similar enough to actually be a dog. However, this is the point at which different people might disagree, and you might change your mind. That is, as similarity gets

lower, there is no clear answer as to whether the item is or isn't in the category. Furthermore, as the item becomes more similar to other categories, the chance increases that it will be seen as an atypical member of that other category. You might decide that your friend's exotic, hairless pet is a weird cat instead of a weird dog.

In short, typicality is a graded phenomenon, in which items can be extremely typical (close to the prototype), moderately typical (fairly close), atypical (not close), and finally borderline category members (things that are about equally distant from two different prototypes). This sort of description summarizes the range of typicality that is often seen in categories. Researchers of all theoretical persuasions talk about category prototypes, borderline cases, and typical and less typical items in this way. However, this way of thinking about typicality should not make us assume that people actually represent the category by a single prototype, like Posner and Keele's original dot pattern, or the best possible dog. There are a number of other ways that they could be representing this information, which I will discuss in the next section and next chapter.

Theoretical Explanations of Prototype Phenomena

What Makes Items Typical and Atypical

What makes something typical or atypical? Why is penguin an atypical bird but sparrow a typical bird? Why should a desk chair be typical but a rocking chair less typical? One possible answer is simple frequency. In North America and Europe (where most of the research on this topic has been done), penguins are seldom if ever seen, whereas sparrows are often seen; there are a lot more desk chairs than rocking chairs, perhaps. This is one part of the answer, but things are not so simple. For example, there are some quite frequent items that are still thought to be atypical of their category. For example, chickens are a very frequently talked-about kind of bird, but they are not considered as typical as some infrequently encountered and discussed birds, such as the oriole or catbird (Rosch 1975) (e.g., I cannot say for certain that I have ever seen an oriole or catbird, but I see chickens fairly often). Similarly, handball is a much more typical sport than car racing (Rosch 1975), even though racing is more popular, is reported in the media, and so on. Mervis et al. (1976) found that simple frequency of an item's name did not predict its typicality. In fact, in the eight common categories they examined (birds, tool, fruit, etc.), none of the correlations of name frequency with typicality was reliably different from 0.

The best answer to the question of what makes something typical is related to the frequency of the item but is a bit more complex: Rosch and Mervis (1975) argued

that items are typical when they have high *family resemblance* with members of the category: "... members of a category come to be viewed as prototypical of the category as a whole in proportion to the extent to which they bear a family resemblance to (have attributes which overlap those of) other members of the category. Conversely, items viewed as most prototypical of one category will be those with least family resemblance to or membership in other categories" (p. 575).[3] That is, typical items (1) tend to have the properties of other category members but (2) tend not to have properties of category nonmembers. For example, an oriole is of the same size as many birds, perches in trees, flies (common bird behaviors), has two wings (universal among birds), and so on. Therefore, even though orioles are not very frequent in my experience, I can recognize them as typical birds because they have frequent bird properties. In contrast, chickens, which I see much more often, are large for birds, white (not a very usual color), do not fly (not common among birds), are eaten (not common), and so on. So, even though chickens themselves are fairly common, their *properties* are not very common as properties for birds, and so they are atypical birds.

Rosch and Mervis (1975) supported this claim in a number of ways. Because their study is so important, I will discuss it in some detail. First, they examined a number of different natural categories[4] to see if they could predict which items would be typical or atypical. Rosch and Mervis selected twenty members of six general categories. Two examples are shown in table 2.1. They had subjects rate each item for how typical it was of its category, and that is reflected in table 2.1 in the order the items are listed (e.g., chair is the most typical furniture, and telephone the least typical). Then Rosch and Mervis had new subjects list the attributes of each of the items. The question was whether typical items would have more of the common attributes than would atypical items. However, there is a problem with just listing attributes, which is that people are not always very good at spontaneously producing attributes for items. If you were to look at such data, you would see that some subjects have produced an attribute for an item, but others have not. Is this because the latter just forgot, or weren't as careful in doing the task? In some cases, that seems likely. Furthermore, it is often the case that attributes listed for one item seem to be equally true for another item, for which it was not listed. For example, people might list "has four legs" for chair but not for bed. Because of this kind of problem (see Tversky and Hemenway 1984 for discussion), Rosch and Mervis asked judges to go through the list of attributes and to decide whether the attributes were true of all the items. So, if some subjects listed "has four legs" for chairs, the judges would decide whether it was also true for the other items of furniture. The judges also elimi-

Table 2.1.
Examples of items studied by Rosch and Mervis (1975), ordered by typicality.

Furniture	Fruit
chair	orange
sofa	apple
table	banana
dresser	peach
desk	pear
bed	apricot
bookcase	plum
footstool	grapes
lamp	strawberry
piano	grapefruit
cushion	pineapple
mirror	blueberry
rug	lemon
radio	watermelon
stove	honeydew
clock	pomegranate
picture	date
closet	coconut
vase	tomato
telephone	olive

nated any features that were clearly and obviously incorrect. This *judge-amending process* is now a standard technique for processing attribute-listing data.

Finally, Rosch and Mervis weighted each attribute by how many items it occurred in. If "has four legs" occurred in ten examples of furniture, it would have a weight of 10; if "is soft" occurred in only two examples, it would receive a weight of 2. Finally, they added up these scores for each feature listed for an item. So, if chair had eighteen features listed, its score would be the sum of the weights of those eighteen features. This technique results in the items with the most common features in the category having the highest scores. Rosch and Mervis found that this feature score was highly predictive of typicality (correlations for individual categories ranged from .84 to .91). That is, the items that were most typical had features that were very common in the category. The less typical items had different features. This result has been replicated by Malt and Smith (1984). Rosch and Mervis illustrated this in a clever way by looking at the five most typical items and five least typical items in each category and counting how many features they each had in common.

They found that the five most typical examples of furniture (chair, sofa, table, dresser, and desk) had thirteen attributes in common. In contrast, the five least typical examples (clock, picture, closet, vase, and telephone) had only two attributes in common. For fruit, the five most typical items had sixteen attributes in common, but the least typical ones had absolutely no attributes in common.

This result gives evidence for the first part of the family-resemblance hypothesis, but what about the second part, that typical items will not have features that are found in other categories? This turns out to be much more difficult to test, because one would have to find out the attributes not just of the target categories (say, fruit and furniture), but also of all other related categories, so that one could see which items have features from those categories. For example, if olives are less typical fruit because they are salty, we need to know if saltiness is a typical attribute of other categories, and so we would need to get feature listings of other categories like vegetables, meats, desserts, grains, and so on, which would be an enormous amount of work. (If I have not made it clear yet, attribute listings are very time-consuming and labor-intensive to collect, as individual subjects can list features of only so many concepts, all of which then must be transcribed, collated, and judge-amended by new subjects.) So, this aspect of the hypothesis is not so easy to test. Nonetheless, Rosch and Mervis were able to test it for two more specific categories (chair and car, Experiment 4), and they found evidence for this hypothesis too. That is, the more often an item had features from a different category, the less typical it was (correlations of $-.67$ and $-.86$).

Both of these studies are correlational. Like most studies of natural categories, the underlying variables that were thought to be responsible for the phenomenon (typicality) were simply observed in the items—the researchers did not manipulate them. As is well known, this leads to the correlation-causation problem, in which there could be some other variable that was not directly measured but was actually responsible for the typicality. Perhaps the most typical items were more familiar, or pleasant, or ... (add your favorite confounding variable). Thus, Rosch and Mervis (1975) also performed experimental studies of family resemblance that were not subject to this problem. They used lists of alphanumeric strings for items, such as GKNTJ and 8SJ3G. Each string was an exemplar, and subjects had to learn to place each exemplar into one of two categories. The exemplars were constructed so that some items had more common features (within the category) than others. Rosch and Mervis found that the items with higher family-resemblance scores were learned sooner and were rated as more typical than were items with lower scores. For these

artificial stimuli, there is little question of other variables that might explain away the results for natural stimuli. In a final experiment, Rosch and Mervis varied the amount of overlap of an item's features with the contrast category. So, one item had only features that occurred in Category 1; another item had four features that occurred in Category 1, and one that occurred in Category 2; a different item had three features that occurred in Category 1, and two that occurred in Category 2; and so on. They found that the items with greater overlap with the other category were harder to learn and were rated as much less typical after the learning phase.

In summary, Rosch and Mervis's (1975) study provided evidence for their family-resemblance hypothesis that items are typical (1) if they have features common in their category and (2) do not have features common to other categories. Unfortunately, there is some confusion in the field over the term "family resemblance." I have been using it here as indicating both parts of Rosch and Mervis's hypothesis. However, other writers sometimes refer to "family resemblance" as being only the within-category component (1 above).[5] In fact, the between-category aspect is often ignored in descriptions of this view. It is important to note, though, that whatever name one uses, it is *both* having features common within a category and the lack of features overlapping other categories that determine typicality according to this view, and Rosch and Mervis provided support for both variables influencing typicality.

This view of what makes items typical has stood the test of time well. The major addition to it was made by Barsalou (1985), who did a more complex study of what determines typicality in natural categories. Using most of the same categories that Rosch and Mervis did, such as vehicles, clothing, birds, and fruit, Barsalou measured three variables. *Central tendency* was, roughly speaking, the family-resemblance idea of Rosch and Mervis, though only including within-category resemblance. The items that were most similar to other items in the category had the highest central tendency scores. *Frequency of instantiation* was the frequency with which an item occurred as a member of the category, as assessed by subjects' ratings. Barsalou felt that the simple frequency of an item (e.g., chicken vs. lark) was probably not as important as how often an item was thought of as being a member of the category, and frequency of instantiation measured this. *Ideals* was the degree to which each item fit the primary goal of each category. Here, Barsalou selected, based on his own intuitions, a dimension that seemed critical to each category. For example, for the category of vehicle, the dimension was how efficient it was for transportation.

Subjects rated the items on these dimensions and also rated the typicality of each item to its category. The results, then, were the correlations between these different measures and each item's typicality.

Barsalou's results provided evidence for all three variables. Indeed, when he statistically controlled for any two of the variables (using a partial correlation), the remaining variable was still significant. Strongest was the central tendency measure. The items that were most similar to other category members were most typical—the partial correlation was .71. This is not surprising, given Rosch and Mervis's results. More surprising was the reliable effect of frequency of instantiation (partial correlation of .36). Consistent with Mervis et al. (1976), Barsalou found that the pure frequency of any item (e.g., the familiarity of chicken and lark) did not predict typicality. However, the frequency with which an item occurred as a category member did predict typicality. This suggests that thinking of an item as being a category member increases its typicality, perhaps through an associative learning process (though see below).

Finally, the effect of ideals was also significant (partial correlation of .45). For example, the most typical vehicles were the ones that were considered efficient means of transportation, and the most typical fruit were those that people liked (to eat). This effect is not very consistent with the Rosch and Mervis proposal, since it does not have to do with the relation of the item to other category members, but with the relation of the item to some more abstract function or goal. Barsalou (1985) presented the results for each category separately, and an examination of the results shows that ideals were quite effective predictors for artifact categories like vehicles, clothing and weapons, but less so for natural kinds like birds, fruit and vegetables.[6] As Barsalou himself notes, since he only used one ideal per category, and since he derived them based on his own intuitions, it is possible that ideals are even more important than his results reveal. If more comprehensive testing for the best ideals had been done, they might have accounted for even more of the typicality ratings. Indeed, a recent study by Lynch, Coley, and Medin (2000) found that tree experts' judgments of typicality were predicted by ideal dimensions of "weediness" and height, but not by the similarity of the trees. Weediness refers to an aesthetic and functional dimension of how useful trees are for landscaping. Strong, healthy, good-looking trees are typical based on this factor. The importance of height is not entirely clear; it may also be aesthetic, or it may be related to family resemblance, in that trees are taller than most plants, and so tall trees are more typical. Surprisingly, the variable of centrality—that is, similarity to other trees (Rosch and Mervis's 1975 first factor)—did not predict typicality above and beyond these two factors.

These results suggest the need for further exploring how important ideals are relative to family resemblance in general.

The importance of ideals in determining typicality is that they suggest that the category's place in some broader knowledge structure could be important. That is, people don't just learn that tools are things like hammers, saws, and planes, but they also use more general knowledge about what tools are used for and why we have them in order to represent the category. Category members that best serve those functions may be seen as more typical, above and beyond the specific features they share with other members.

As mentioned above, correlational studies of this sort, although important to carry out, can have the problem of unclear causal connections. In this case, the frequency of instantiation variable has a problem of interpreting the directionality of the causation. Is it that frequently thinking of something as being a category member makes it more typical? Or is it that more typical things are likely to be thought of as being category members? Indeed, the second possibility seems more likely to me. For example, Barsalou found that typical items are more likely to be generated first when people think of a category. When people think of examples of weapons, for example, they hardly ever start with spears or guard dogs; they are more likely to start with typical items like guns and knives. Frequency of instantiation, then, could well be an effect of the typicality of items, rather than vice versa. Similarly, the effect of ideals might also have come after the category's typicality was determined. That is, perhaps the ideal of vehicles came about (or occurred to Barsalou) as a result of thinking about typical vehicles and what their purposes were. It is not clear which is the chicken and which is the egg.

Barsalou addressed this problem in an experimental study of the importance of ideals. He taught subjects new categories of kinds of people. Subjects in one category tended to jog (though at different frequencies), and subjects in another tended to read the newspaper (at different frequencies). Barsalou told subjects that the two categories were either physical education teachers and current events teachers, or else that they were teachers of computer programming languages Q or Z. His idea was that the joggers would tend to be perceived as fulfilling an ideal dimension (physical fitness) when the category was physical education teachers but not when the category was teachers of language Q; similarly, the newspaper readers would fit an ideal (being informed) when they were current events teachers but not when they were teachers of language Z. And indeed, amount of jogging influenced typicality ratings for the categories described as physical education teachers, but not for the category labeled teachers of language Q even though both categories had the exact

same learning items. That is, the family-resemblance structures of the items were identical—what varied was which ideal was evoked.

This experiment, then, confirms part of the correlational study that Barsalou (1985) reported for natural categories. Ideals are important for determining typicality above and beyond any effects of family resemblance. A number of later studies have replicated this kind of result, and are discussed in chapter 6 (Murphy and Allopenna 1994; Spalding and Murphy 1996; Wattenmaker et al. 1986; Wisniewski 1995). Studies of experts have found even more evidence of the use of ideals (Medin, Lynch, Coley, and Atran 1997; Proffitt, Coley, and Medin 2000).

These studies do not show that family resemblance is not a determinant of typicality, but that there are also other determinants that Rosch and Mervis (1975) would not have predicted. Thus, I am not suggesting that their view is incorrect but that it is only a partial story, for some categories at least. They are clearly correct, however, in saying that frequency in and of itself does not account for typicality to any great degree (see also Malt and Smith 1982). Whether frequency of instantiation is an important variable is less clear. Barsalou's results do give evidence for its being related to typicality. However, since this variable has not been experimentally tested, the result is subject to the counterargument raised above, that it is a function of typicality rather than vice versa. This, therefore, is still an open issue.

End of the Classical View?

The classical view appears only very sporadically after this point in the book. To a considerable degree, it has simply ceased to be a serious contender in the psychology of concepts. The reasons, described at length above, can be summarized as follows. First, it has been extremely difficult to find definitions for most natural categories, and even harder to find definitions that are plausible psychological representations that people of all ages would be likely to use. Second, the phenomena of typicality and unclear membership are both unpredicted by the classical view. It must be augmented with other assumptions—which are exactly the assumptions of the nonclassical theories—to explain these things. Third, the existence of intransitive category decisions (car seats are chairs; chairs are furniture; but car seats are not furniture) is very difficult to explain on the classical view. The classical view has not predicted many other phenomena of considerable interest in the field, which we will be discussing in later chapters, such as exemplar effects, base rate neglect, the existence of a basic level of categorization, the order in which children acquire words, and so on. In some cases, it is very difficult to see how to adapt this view to be consistent with those effects.

In summary, the classical core has very little to do, as most of the interesting and robust effects in the field cannot be explained by cores but must refer to the characteristic features. This is true not only for speeded judgments, but even for careful categorizations done without time constraints.

In spite of all this, there have been a number of attempts to revive the classical view in theoretical articles, sometimes written by philosophers rather than by psychologists (Armstrong et al. 1983; Rey 1983; Sutcliffe 1993). This is a bit difficult to explain, because the view simply does not predict the main phenomena in the field, even if it can be augmented and changed in various ways to be more consistent with them (as discussed in the core-plus-identification procedure idea). Why are writers so interested in saving it? To some degree, I believe that it is for historical reasons. After all, this is a view that has a long history in psychology, and in fact has been part of Western thinking since Aristotle. (Aristotle!) If this were a theory that you or I had made up, it would not have continued to receive this attention after the first 10 or 20 experiments showing it was wrong. But our theory would not have this history behind it. Another reason is that there is a beauty and simplicity in the classical view that succeeding theories do not have. It is consistent with the law of excluded middle and other traditional rules of logic beloved of philosophers. To be able to identify concepts through definitions of sufficient and necessary properties is an elegant way of categorizing the world, and it avoids a lot of sloppiness that comes about through prototype concepts (like the intransitive category decisions, and unclear members). Unfortunately, the world appears to be a sloppy place.

A final reason that these revivals of the classical view are attempted is, I believe, because the proponents usually do not attempt to explain a wide range of empirical data. The emphasis on philosophical concerns is related to this. Most of these writers are not starting from data and trying to explain them but are instead starting from a theoretical position that requires classical concepts and then are trying to address the psychological criticisms of it.[7] In fact, much of the support such writers give *for* the classical view is simply criticism of the evidence *against* it. For example, it has been argued that typicality is not necessarily inconsistent with a classical concept for various reasons, or that our inability to think of definitions is not really a problem. However, even if these arguments were true (and I don't think they are), this is a far cry from actually explaining the evidence. A theory should not be held just because the criticisms of it can be argued against—the theory must itself provide a compelling account of the data. People who held the classical view in 1972 certainly did not predict the results of Rips et al. (1973), Rosch (1975), Rosch and Mervis (1975), Hampton (1979, 1988b, or 1995), or any of the other experiments that helped to

overturn it. It was only after such data were well established that classical theorists proposed reasons for why these results might not pose problems for an updated classical view.

As I mentioned earlier, there is no specific theory of concept representation that is based on the classical view at the time of this writing, even though there are a number of writers who profess to believe in this view. The most popular theories of concepts are based on prototype or exemplar theories that are strongly unclassical. Until there is a more concrete proposal that is "classical" and that can positively explain a wide variety of evidence of typicality effects (rather than simply criticize the arguments against it), we must conclude that this theory is not a contender. Thus, although it pops up again from time to time, I will not be evaluating it in detail in the remainder of this book.